

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 14:43:51 ; Search time 85.35 Seconds
(without alignments)
4990.375 Million cell updates/sec

Title: US-09-710-339-1

Perfect score: 1734
Sequence: 1 tcatcatcaagcttccttc.....aaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCRUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	375.4	21.6	1404	US-08-204-656B-7	Sequence 7, Appl
2	375.4	21.6	1404	US-08-470-702-4	Sequence 4, Appl
3	375.4	21.6	1404	US-08-467-831-4	Sequence 4, Appl
4	373.8	21.6	1404	US-08-204-656B-1	Sequence 1, Appl
5	373.8	21.6	1404	US-08-204-656B-3	Sequence 3, Appl
6	373.8	21.6	1404	US-08-204-656B-5	Sequence 5, Appl
7	373.8	21.6	1404	US-08-470-702-1	Sequence 1, Appl
8	373.8	21.6	1404	US-08-470-702-2	Sequence 2, Appl
9	373.8	21.6	1404	US-08-470-702-3	Sequence 3, Appl
10	373.8	21.6	1404	US-08-467-831-1	Sequence 1, Appl
11	373.8	21.6	1404	US-08-467-831-2	Sequence 2, Appl
12	373.8	21.6	1404	US-08-467-831-3	Sequence 3, Appl
13	68.6	4.0	93	US-09-189-060B-63	Sequence 63, Appl
14	68.4	3.9	919	US-09-189-060B-67	Sequence 67, Appl
15	68.4	3.9	922	US-09-189-060B-65	Sequence 65, Appl
16	68.4	3.9	922	US-09-189-060B-73	Sequence 73, Appl
17	68.4	3.9	924	US-09-189-060B-69	Sequence 69, Appl
18	63.4	3.7	912	US-09-189-060B-71	Sequence 71, Appl
19	58	3.3	2582	US-08-816-105A-2	Sequence 2, Appl
20	55.4	3.3	5163	US-08-700-651-1	Sequence 1, Appl
21	55.4	3.2	5163	US-08-928-361B-4	Sequence 4, Appl
22	55.4	3.2	5318	US-08-700-651-2	Sequence 2, Appl
23	55.4	3.2	5318	US-08-928-361B-3	Sequence 3, Appl
24	51.4	3.0	5511	US-08-928-361B-2	Sequence 2, Appl
25	51.4	3.0	7334	US-08-928-361B-1	Sequence 1, Appl
26	50.6	2.9	2160	US-09-386-607-1	Sequence 1, Appl
27	49.2	2.8	1295	US-08-433-854-3	Sequence 3, Appl

28	49.2	2.8	1295	US-08-174-745A-3	Sequence 3, Appl
29	49.2	2.8	1295	US-08-195-947-3	Sequence 3, Appl
30	49.2	2.8	1295	US-08-433-885-3	Sequence 3, Appl
31	49.2	2.8	1295	US-08-433-908B-3	Sequence 3, Appl
32	49.2	2.8	1295	US-08-410-614-3	Sequence 3, Appl
33	48.4	2.8	390	US-09-197-649-7	Sequence 7, Appl
34	44.8	2.6	1100	US-07-861-458C-4	Sequence 4, Appl
35	44.8	2.6	2781	US-08-140-008A-1	Sequence 1, Appl
36	43.8	2.5	2061	US-08-204-656B-9	Sequence 9, Appl
37	43.8	2.5	2061	US-08-470-702-5	Sequence 5, Appl
38	43.8	2.5	2061	US-08-467-831-5	Sequence 5, Appl
39	43.2	2.5	5580	US-08-814-052-10	Sequence 10, Appl
40	43.2	2.5	5580	US-08-812-829-10	Sequence 10, Appl
41	43.2	2.5	5679	US-08-814-052-9	Sequence 9, Appl
42	43.2	2.5	5679	US-08-812-829-9	Sequence 9, Appl
43	43.2	2.5	5697	US-08-814-052-11	Sequence 11, Appl
44	43.2	2.5	5697	US-08-812-829-11	Sequence 11, Appl
45	43	2.5	2209	US-08-514-014-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-204-656B-7
; Sequence 7, Application US/08204656B
; Patent No. 5538882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuo
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; APPLICANT: Honda, Koichi
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 810 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,656B
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: A35
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: Derived from plasmid pSF11 (Agric. Biol. Chem.
; FEATURE:

SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-4

Query Match 21.6%; Score 375.4; DB 1; Length 1404;
Best Local Similarity 55.6%; Pred. No. 5.6e-98;
Matches 752; Conservative 0; Mismatches 586; Indels 15; Gaps 1;

QY 121 ctggggactgagtcgcaatccatcttattctcctcccaagatcgattggcaagcg 180
DB 11 CTGTAATGAGATCAGCTATTTATCAAAATGTCATGACAGATTGGCTAGAACCG 70
QY 181 atggctcagcactgcaactgttaatactgcgagtcagaaatactggtggaacatg 240
DB 71 ATGGGATACAGAGTCTTCCTGTAAACAGAGATAGACTTACTGTGGGTCTTCC 130
QY 241 agggcatatcgacaagttgactatatccaggaatggcgcttcacagccatcgatca 300
DB 131 AAGGCATCATAAAGATGGATTACATCAAAAGATATGGGCTTACATGCTATTTGATTT 190
QY 301 ccccgatcaagcccaagtcgcccagaccagcatatgagagatgcatcactgact 360
DB 191 CTCGATGTTGAAATCTCCGATTAACACAGATATGGTTATGCTTATCATGATACT 250
QY 361 ggcagcagatatatactctctgacgaaactacgacgcaactgcagatgacttgaagcgc 420
DB 251 GGATGAAGACATATATCAAAATTAATGAACCTTGGTACTGCGATGATTTGAAGCTT 310
QY 421 tctctcgccctcatgagagggggtatctctatggtcgatggttgctaacata 480
DB 311 TGGCAAGAAATTCACGATGCTATATGTTAATGGTGGATATGCTTACCAACATT 370
QY 481 tgggctatgagggcggtgagtcacagtcagatcaggtgtttaaaccgttcctcc 540
DB 371 ACGGATGATGGCAGTGGAGATAGTATGATCTCAGATTAACCCCGTTCAACGACC 430
QY 541 aagactatccaccggtctgttcaatcaaaactaagaatcagactcaggttgag 600
DB 431 AAAAGTACTTCATTAAGTACTGTCTTATTTCAAACTATGATGACCAAGCTCAGGTTCAA 490
QY 601 attgctgctagagagataaacctgtctccttgctgactcgcataccaagaagtgtg 660
DB 491 GTTGTGGGAAGGTGACTCTTCAGTTGATTAACAGATTTGAGAAACGGAAGATAGCGACG 550
QY 661 tcaagaatgaatgtagcagtcggttggagatcatgtagtgcgaactcaactcagagcc 720
DB 551 TGGCTCAGTTTTCATTTCTGGGTTAAAGATTTTGTGGCAATTTACTCAATTTATGTGTT 610
QY 721 tccglatcgacacagtaaacacgctccagaagactctcgtgcccgggtacaacaagccg 780
DB 611 TAAGATTTGATAGTCTTAACATGTGACCAAGGCTTTTCCCGATTTTGTGTAGTCAT 670
QY 781 caggcggtactgtatcgcgaggtgctcgagcaggtgtagccggctcactgttcctacc 840
DB 671 CTGGAGTTTACTCAGTACGCGAAGTTTTCAGAGGACCCAGCTTATTAAGGCCATACC 730
QY 841 agaagcgtatgagcggtactgaactatcccatctatccactcctcctaagccctca 900
DB 731 AAAATTAATTCAGAGGGGTATGATATTTCAATTTACTACCAACCGAGATTTTTTTA 790
QY 901 agtcaacctccggcagcaatgtagcagcctctacaacatgatacaacgcgtcaaatccgact 960
DB 791 AAACATGATTTCAAGTTCACAGTGAAGTGAAGTCAAAATGATTTTCAAGCGTGTCTTCAAGTT 850
QY 961 gtccagactcaaacactcctgggacatctcgtagaacaacagacaaccacgggttcgtt 1020

DB 851 GTTGGATCCAACTTTGTTGACAAACTTTGAGAAAATCAGATATGAAGTTGCTT 910
QY 1021 cttaaccaagacatagccctcgccaagaacgtcgcaacttcaatcctcacaagc 1080
DB 911 CAATGACCAAGCCAAAGTTGATTTCTAATGCTATTCATTTGCTTTGGGTGATG 970
QY 1081 gaatcccatctatcgccggcccaagaacagcaactacgacgaggaacgaacgagcga 1140
DB 971 GATTCCTGCTATTTACTATGAGCAAGAACAGGCTTGAGCGGAAAGTGACCCAAACA 1030
QY 1141 accggaagcaactggtcctcggtctaccgacagcagcgagctgtataagttaatg 1200
DB 1031 ACAGAGAGGCTTGTGTTATCCGCTTACAAACAAAGAGATGACTTATTAACAGCTATTG 1090
QY 1201 ccccgagcaagcaatccgaaactatgcatatagcaaatgacagatcgtgacttaca 1260
DB 1091 CCAAGCTATATGCTGCGAAGAAAGCCGCGCTTATCAACATCAAGCTATGCGACCTGCG 1150
QY 1261 agaatgccccatctacaagaacgacaacagatcgccatgccaagggcacagatgggt 1320
DB 1151 AGCTTCTGTGATCTTTCAAAATGACCATGTTATGCAACAAAAGAGCGACGTTGTTT 1210
QY 1321 cgcagatcgtgatactctgtccaagaaggtgcttcggtgattcgtataccctcct 1380
DB 1211 CTG-----TTTTCACAAACCTTGTTCCAGCGGTTCTTCTGTGATGTGACTA 1255
QY 1381 tgaatggtcggtgtacacaagccggccgaatgtagcgaaggtcattgctgacagccg 1440
DB 1256 TTTCCAAACACAGGTTACAGTCCGGTGAGGATTTGGTAAAGTTTGCATCAGTACTG 1315
QY 1441 tgaaggttggttcgaltggaatgltgcctgttc 1473
DB 1316 TTAGCGGACAGCTCTGACTTACAGTTTCTATGCC 1348

RESULT 3
US-08-467-831-4
Sequence 4, Application US/08467831
Patent No. 5635378
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Galehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,831
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000


```

STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: Derived from plasmid pSF1 (Agric. Biol. Chem.
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1404
OTHER INFORMATION: /note= "Nucleotides 1-1404
correspond to nucleotides 79-1482 of the Saccharomyces
OTHER INFORMATION: fibuligera - amylase structural gene"
US-08-204-656B-5

Query Match      21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 1.6e-97;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

```

```

QY 541 aagactactccaccgcgtctcttccatcaaaactatgaatcagctcaggtttag 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 431 AAAGTACTTCCATTACTACTGCTTATTCAACATGATGATCCAAAGCTCAGTTCAA 490
QY 601 attcgtcgttagagataaactgtctcctgtcgtatcgtatcaccagaagttag 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 491 GTTGCGGGAAGGTGACTCTTCAGTTCATTTACCAATTTGAGAACGGAAGATAGCGAG 550
QY 661 tcaagaatgatgtgtcagactgggttggtatcattgtatcgaactcctcagcgcc 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 551 TGCCCTCAGTTTCAATTTCTTGGGTTAAAGATTTTGTGGCAATTTCTCAATTTGATGTT 610
QY 721 tccgatacgaacagtaaaacacgctccagaaggactctcggcggtatacaaaagccg 780
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 611 TAAGAAATGATAGTGGCAAAACATGTGGACCAAGCTTTTCCCGATTTGTTAGTCAT 670
QY 781 caggcgtgtactgtatcgcgcgaggtgtcgcagcgtgatccggcctactgtccctacc 840
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 671 CTGGAGTTTACTCAGTAGGCGAAGTTTCCAAAGGAACCCAGCTTATACATGCGCATACC 730
QY 841 agaacgtcatcagcggcggtactgaactatccattactatccactcctoaagccttca 900
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 AAATTTACATTTCCAGGGGTTAGTAATTTATCCATTGTACTCCACACAGAGATTTTAA 790
QY 901 agtcaacctccgcgagcatatgacgacctctacaacatgatcaaacocgtcaatccgact 960
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 AAATCTCATGATTCAAGTTCCAGTAGTTGACTCAAAATGATTCAAGGCTTGCCTCAGTT 850
QY 961 gtccagactcaaacctcctgtggcactgtcgtgagaacccgcgaaccccggttcgt 1020
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 851 GTTCGATTCACAACTTTGTGCAAACTTTGTAGAAATTCACAGTAATGAAAGGTGCTTT 910
QY 1021 cttaacccaagacatagccctcgcgaagacgctcgaagatcatcatcctcaacgacg 1080
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 911 CAATGACCAAGGACCAAAAGTTGATTCTTAATGCTATTTGCAATTTGGGTGATG 970
QY 1081 gaatcccatcatcctacgcgcgcgaagaacgacactcgcgcgcgaagaccccgga 1140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 971 GTATTCTCGTATTACTATGATGACAAAGAACAGCTTGAGCGGAAAAAGTGACCAAAACA 1030
QY 1141 accgcgaagcaactgtgtctcgcggtacccgacccgagcagctgtacaagttaattg 1200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1031 ACAGAGAGGCGCTTGTTGGTTATCCGCTACAAACAAAGAGATGATATTACAGCTCTTG 1090
QY 1201 cctccgcgaacgcaatccgcgaatcgtcattgaagaatgacagatcgtcgtactaca 1260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1091 CCAAGGCTAATGCTGCCAGAAAGCGCCGTTTATCAAGACTCAAGCTATCCACCTCGC 1150
QY 1261 agaactggccaactctacaagaacgacacacgactcgcgtatcgcgaagggcagaatgg 1320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1151 AGCTTTCTGTGATCTTTTCAAAATGACCATGTTATGCAAAAAGAGGCGAGCTTGTT 1210
QY 1321 gcgagatcgtgtactatctgtccaaagaaggtgtctcgggtatcgtatccctcct 1380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1211 CTG-----TTTTCAAACACTTTGGTTCCAGCGGTTCTGTGATGACTA 1255
QY 1381 tgaagtgtgcggttaccacagccgcgcagcaatltgacggaaggtcattgtgtcagacgcg 1440
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1256 TTTCCAAACAGAGTTTACAGTTCCGGGTGAGGATTTGTTGAAGTTTTCATGACGACTG 1315
QY 1441 tgaaggtgtgtcgtatggaatgtgctgttc 1473
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1316 TTAGCGGAGCTGTGACTTACAAAGTTTCTATCC 1348

```

```

RESULT 7
US-08-470-702-1
; Sequence 1, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO

```


APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-2

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 1.6e-97;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 ctggagactcgagatcgcaatcattatctctcctcagatcgatttggaagagcg 180
DB 11 ctgattaaatggagatcagatctattttatcattgtctcactcagatttgctaaacgg 70
QY 181 atgggctgagactgagacttgtaactcgatcagaatactgctggtggaacatgac 240
DB 71 atggtatcacaactgcttctctgtaacacagaaatgactttactgtgtgcttttc 130
QY 241 agggcactcgacaagtggagactatcaccaggaatgggcttcacagccatctgata 300
DB 131 aagcgtatcattaaagagttgatttattcattcattgagcttttctgatttgat 190
QY 301 ccccgcttaacgcccagcccgacacacacacacacacacacacacacacacacacac 360
DB 191 ctccagttgttgaataacattccgattacacacacacacacacacacacacacacac 250
QY 361 ggcagcagatatactctctgaacgaataactcgcacatgacttgaagcg 420
DB 251 ggatgaagaacatattacaaatattgaataattgtgtactgtgtgattgaactgt 310
QY 421 tctctcgccctcatgagagggatgtatcttatgttcgactgtgtgtgcttaaccata 480

DB 311 tggcacaaagattgcacgactcgatattgtttatggtgacgatttgcataacacatt 370
QY 481 tggcctatgatagagaggttagctcagctgattacagtggttttaacgctcaagtc 540
DB 371 acggcgtatagcagctgagatgattatgatttactcagagatgacccggttcaacgac 430
QY 541 aagaactcacaacccgctctgttcaatcaaaactatgaatgaatcagactcagttgag 600
DB 431 aaaaacttccatattcactgcttattttcaaaactatgattgacacacacgctcagttcaa 490
QY 601 attgtgcttagagataaacaactgtctctcctcgtatcgtatccacacagatgag 660
DB 491 gttcgtggagagctgactcttccatttccatttccatttgaacaggaattgacgacg 550
QY 661 tcaaatatgatgtacagactggtggtggtatcttgatcgaactcactcactgacgac 720
DB 551 tggcctcagttttcaatttcttggtttaaagatttttggcatttactcatttgaattg 610
QY 721 tccgtaacagtaaaacacgctccagaagactctgcccgggtacaaagacg 780
DB 611 taagaattgtactgcttaaaacatgacacacgcttttcccgattgtttagtgcat 670
QY 781 caggcgtgactgtatcgcgcgaggtgctgacggtgataccgctcactgtccctacc 840
DB 671 ctggagtttactcagtagccgaagttttccaaagagacccagcttttaccattgacatcc 730
QY 841 agaacgtcatgagcggcgtactgaactatccattactatccactcactcactcaagctca 900
DB 731 aaatttcatattccaggggttagtattttatccattgattaccacacagagatttttta 790
QY 901 agtaacctccgycagatagagacacttacaacatgatcaacacgctcaaatccgact 960
DB 791 aaactactgatttcaagttccagtagttgactcaaatgattttcaacgcttccagtt 850
QY 961 gtccagactcaaacctctggtgacatctgctgagagacacgacacacacacacacac 1020
DB 851 gttcgatccacttgtttgacaaactttgttgaataatcgaatttgaagagtttgctt 910
QY 1021 cttaacacacagacatagccctcgccaagaacgctgcagatcattacatcctcaacagc 1080
DB 911 caatgaccacagacacaaagtttgaattttctaatgctatttgcatttgggtgattg 970
QY 1081 gaatcccatcattcagccggtccgaagaacgactcgcgcgcggaagacaccccgca 1140
DB 971 gtatttccgtcatttactatgacacaaagacgcttgagcgaacacaaagtgacccaa 1030
QY 1141 accggaagacacactgctcgcggtcctacccgacacgacgactgttacaattatg 1200
DB 1031 acagagagcccttgggttattccgcttacaacaaagagatgactatttacaactcatt 1090
QY 1201 cctcgcgaacgcaatccggaactatgcatltagcaaatagatcacagatctcgtaactaca 1260
DB 1091 ccaaaagctaatgctgccaagaaccccggtttttcaaacctcaagctatgaccacgctgc 1150
QY 1261 agaactgcccactacacaagaacacacacacacacacacacacacacacacacacac 1320
DB 1151 agcttttctgtatcttttcaaatgacacatttattgacacaaagagacgacgctgttt 1210
QY 1321 cgcagatcgtgactatctctgccaagaaggtgctcgtggtgattatccctcct 1380
DB 1211 ctg-----ttttcaacaaacttttgcctcagcggttcttgcattgtgacta 1255
QY 1381 tgaatgtgctggttacaacagcgcgcgacacattgacgaggttcaattgctcgaacg 1440
DB 1256 tttccaaacacaggttaccggttccggtgagattgagatttgaacatgacagttactg 1315
QY 1441 tgaagctgtgttcgattggaatgtgctgttc 1473
DB 1316 ttaagcgacgctcgtacttaccagattttctatcc 1348

RESULT 9
US-08-470-702-3


```

RESULT 10
US-08-467-831-1
; Sequence 1, Application US/08467831
; Patent No. 5635378
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,831
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-467-831-1

Query Match      21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 1.6e-97;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

```

```

Qy 361 ggcagcagatatatactctcgaacgaaactacgacgtcgaatgacgtcgaagcgc 420
   || || || || || || || || || || || || || || || || || || || ||
Db 251 GGAATGAAGAAACATATATCAAAATTAATGAACACTTGGTACTGCTGATGATGAAGTCTT 310
Qy 421 tctcttcgcacctcattgagaggggagtgatcctatgctgctgctgctgctgctgct 480
   || || || || || || || || || || || || || || || || || || || ||
Db 311 TGGCACAAGAAATGACAGATCGGATATGTTGTAATGGTGAATATCTTACCAACCAT 370
Qy 481 tggcgtatgagcggagcgtgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 540
   || || || || || || || || || || || || || || || || || || || ||
Db 371 ACGGAGATGATGACATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 430
Qy 541 aagactactccacccgcttctgcttctcattcaaaaactaagaatcagactcaggtttag 600
   || || || || || || || || || || || || || || || || || || || ||
Db 431 AAAAGTACTTCCATTACTACTGCTTATTCAAACTATGATGACCAAGCTCAGGTCAAA 490
Qy 601 attgctgctgagagataaacgtctccttctgctgctgctgctgctgctgctgctgctg 660
   || || || || || || || || || || || || || || || || || || || ||
Db 491 GTTCCTGGGAAGGTGACTCTTCAGTTGCATTTACCAATTTGAGAACGGAAGATAGCGACG 550
Qy 661 tcaagaaatgtagtgcagactggtgagatcaltggtatcgaactactcacttgaagcgc 720
   || || || || || || || || || || || || || || || || || || || ||
Db 551 TGGCCTCAGTTTCAATCTTGGGTTAAAGATTTTGTGGCAATTTCTCAATTTGATGATG 610
Qy 721 tccgtatcgaacaaagtaaaacacgctccagaagactctgcccgggtacacaagccg 780
   || || || || || || || || || || || || || || || || || || || ||
Db 611 TAAGAAATTGATAGTAAACATGATGACCAAGCTTTTCCCGAATTTGTTAGTCAT 670
Qy 781 cagcggtactgtatcggcgaagtgctgacggtatccggctcactgtccctacc 840
   || || || || || || || || || || || || || || || || || || || ||
Db 671 CTGGAGTTTACTCAATGAGGGAAGTTTCCAAAGAGACCCAGCTTTATACATGCCCATACC 730
Qy 841 agaagcgcagcagcgcgtactgaactacatccacttcaactcactcactcactcactc 900
   || || || || || || || || || || || || || || || || || || || ||
Db 731 AAATTTACATTTCCAGGGGTTAGTAATATATCCATTTGACTACCAACACAGATTTTAA 790
Qy 901 agtcaacctccgcagatgaggaagactctcaacaactgatacaacacgctgaactcgcact 960
   || || || || || || || || || || || || || || || || || || || ||
Db 791 AAATCTGATTAAGTTTCCAGTACGATGATGATGATGATGATGATGATGATGATGATGAT 850
Qy 961 gtccagactcaacactcctcgtggcacatcgtcgcgaacacacacacacacacacacac 1020
   || || || || || || || || || || || || || || || || || || || ||
Db 851 GTTGGATCCACTTTGTTGACAAACTTTGTAGAAATTCAGTAATGAAGATGGTGGCTT 910
Qy 1021 cttaacacaacgacatagccctcgcgaagacgctgcgaactcactcactcactcactc 1080
   || || || || || || || || || || || || || || || || || || || ||
Db 911 CAATGACACGAGCACAAGTTGATTTCTAATGCTATTGATTTGCTTTTGGGTGATG 970
Qy 1081 gaatcccatcatctacgcccgcgaagaacacacacacacacacacacacacacacac 1140
   || || || || || || || || || || || || || || || || || || || ||
Db 971 GTATTCCTGCTATTACTATGATGACAAAGAAAGGCTTGAGCGGAAAGATGACCCAAACA 1030
Qy 1141 acccggaagcaacctggtcgcggtcactaccgacacacacacacacacacacacacacac 1200
   || || || || || || || || || || || || || || || || || || || ||
Db 1031 ACAGAGAGCCTTGTGTGTTATCCGGCTACAAACAAAGAGATGATATTACAGTCTATTG 1090
Qy 1201 cctccgcgaacgcaatcccggaactatgccaatagcaaaatagacagatcgtgacctaca 1260
   || || || || || || || || || || || || || || || || || || || ||
Db 1091 CCAAGGTAATGCTGCCAGAAACGCCGCTTTATCAAGACTCAAGCTATGACCACTCGC 1150
Qy 1261 agaactggcccatctacaagaagacacacacacacacacacacacacacacacacacac 1320
   || || || || || || || || || || || || || || || || || || || ||
Db 1151 AGCTTTGCTGATCTTTTCAAAATGACATGTTATTTGCAAAAAGAGAGGACGCTTGT 1210
Qy 1321 cgcagatcgtactactctgtccacaagaagtgctcgtggtatcgtatataccctcct 1380
   || || || || || || || || || || || || || || || || || || || ||
Db 1211 CTG-----TTTTCAAACATCTTGTTGACAGGGTTCTCTGATGATGACTA 1255
Qy 1381 tgaagtgtgaggttacacacacgcccgcagcaatgacggaggtcattggtcgtacagccg 1440
   || || || || || || || || || || || || || || || || || || || ||
Db 1256 TTTCCAAACACAGGTTACAGTTCCGGTGAAGATTTGGTGAAGATTTTGACATGACGACTG 1315
Qy 1441 tgacggtgtgtcgaatgagaaatgtgcctgttc 1473

```




GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 13:40:45 ; Search time 1884.09 Seconds
(without alignments)
12421.763 Million cell updates/sec

Title: US-09-710-339-1

Perfect score: 1734
Sequence: 1 tcacatcaagctctcccttc.....aaaaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gd_estl:*
10: gd_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_iny:*
15: em_gss_pln:*
16: em_gss_vtl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	256.8	14.8	583	AA786071	AA786071 j4h07a1.r
2	240.2	13.9	541	AA787396	AA787396 n3a12a1.r
3	227.2	13.1	561	AA783132	AA783132 c1c1a1.r
4	210.2	12.1	464	AA965792	AA965792 o5e03a1.r
5	189.8	10.9	466	AA781322	AA781322 o5e07a1.r
6	150.8	8.7	347	AA785831	AA785831 h8a04a1.r
7	143.2	8.3	376	AA786814	AA786814 m6c11a1.r
8	140.2	8.1	374	AA785147	AA785147 g5c10a1.r
9	139.4	8.0	382	AA785574	AA785574 g8h01a1.r
10	138.6	8.0	358	AA785574	AA785574 g8h01a1.r
11	133	7.7	389	AA784386	AA784386 x1f04a1.r
12	131.8	7.6	358	AA784386	AA784386 d5d11a1.r
13	131.2	7.6	304	AA787924	AA787924 r2h03a1.r
14	116.2	6.7	405	AA787924	AA787924 r2h03a1.r
15	112.2	6.5	362	AA787426	AA787426 n3c10a1.r
16	111.6	6.4	293	AA785338	AA785338 g8d04a1.r
17	104	6.0	283	AA788570	AA788570 a1f06c9.r

18	103.8	6.0	548	10	BG278743	BG278743 a6h08np.r
19	96.6	5.6	335	9	AI210822	AI210822 10e09a1.r
20	83.2	4.8	450	9	AI210074	AI210074 g9b03a1.r
21	75.6	4.4	876	10	BM400915	BM400915 5009-0-80
22	74.6	4.3	287	10	AA785414	AA785414 g7b08a1.r
23	74.6	4.3	402	12	A2877703	A2877703 #06 3.0 M
24	72.6	4.2	668	9	AU060851	AU060851 AU060851
25	72.6	4.2	677	9	AU060198	AU060198 AU060198
26	71.4	4.1	232	9	AI209785	AI209785 c9f06a1.r
27	69.4	3.5	559	9	AU052395	AU052395 AU052395
28	59.4	3.4	293	9	AA901823	AA901823 NCM9B11T3
29	59.4	3.4	353	9	AI210266	AI210266 h4b03a1.r
30	58.8	3.4	653	9	AU039903	AU039903 AU039903
31	58.8	3.3	457	9	AU034096	AU034096 AU034096
32	57	3.3	457	9	AA787302	AA787302 n1e09a1.r
33	55.2	3.2	500	12	B67199	B67199 CpG0015B Cp
34	54.8	3.2	681	12	CNS02E0D	AI133990 Netradon
35	54.4	3.1	525	9	AI544417	AI544417 SD10962.5
36	53.6	3.1	608	10	BI887904	BI887904 ZF637-1.0
37	53.2	3.1	645	9	AI389106	AI389106 GH20192.5
38	53	3.1	738	12	PT008M15U	AL447629 Parameciu
39	52.8	3.0	649	10	BI641907	BI641907 SD25571.5
40	52.2	3.0	703	12	AO162044	AO162044 mgxb0010F
41	52.2	3.0	895	12	CNS0071A	AL066286 Drosophila
42	51.2	3.0	511	10	BM004465	BM004465 TgESTy28
43	51.2	3.0	922	12	CNS0073W	AL066784 Drosophila
44	51.2	3.0	587	10	BE446309	BE446309 WHE1455_H
45	51	2.9	626	10	BF473952	BF473952 WHE0839_C

ALIGNMENTS

RESULT 1
LOCUS AA786071 583 bp mRNA linear EST 31-JUN-1998
DEFINITION j4h07a1.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone j4h07a1 5', mRNA sequence:

ACCESSION AA786071.1 GI:2846239
VERSION
KEYWORDS
SOURCE Emericella nidulans.
ORGANISM Emericella nidulans.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiiales; Trichocomaceae; Emericella.

REFERENCE 1 (bases 1 to 583)
Kupfer,D., Gray,U., Hausner,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,B.
An Aspergillus nidulans EST Database
Unpublished (1998)
other_ESTs: j4h07a1.f1
Contact: Bruce A. Roe, University of Oklahoma, broeou@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broeou@ou.edu

We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
Seq primer: SK
High quality sequence stop: 526.
Location/Qualifiers
1..583
Location/Qualifiers

FEATURES
source
/organism="Emericella nidulans"
/strain="FGSC A26"
/db_xref="taxon:162425"
/clone="j4h07a1"
/clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
/tissue_type="vegetative mycelia, asexual structures"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

Query Match 14.8%; Score 256.8; DB 9; Length 583;
 Best Local Similarity 65.2%; Pred. No. 1.9e-29;
 Matches 378; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cdna clones to the Fungal
 Genetics Stock Center
 High quality sequence stop: 510.
 Location/Qualifiers
 1..541
 /organism="Emicella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="n3a12a1"
 /clone_11b="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
 XhoI; 5' end of cdna cloned into EcoRI site of pBluescript
 3' end of cdna cloned into XhoI site of pBluescript"

620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cdna clones to the Fungal
 Genetics Stock Center
 High quality sequence stop: 510.
 Location/Qualifiers
 1..541
 /organism="Emicella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="n3a12a1"
 /clone_11b="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
 XhoI; 5' end of cdna cloned into EcoRI site of pBluescript
 3' end of cdna cloned into XhoI site of pBluescript"

Query Match 13.9%; Score 240.2; DB 9; Length 541;
 Best Local Similarity 65.2%; Pred. No. 6.2e-27;
 Matches 353; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cdna clones to the Fungal
 Genetics Stock Center
 High quality sequence stop: 510.
 Location/Qualifiers
 1..541
 /organism="Emicella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="n3a12a1"
 /clone_11b="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
 XhoI; 5' end of cdna cloned into EcoRI site of pBluescript
 3' end of cdna cloned into XhoI site of pBluescript"

RESULT 2
 AA787396 541 bp mRNA linear EST 31-JUL-1998
 LOCUS n3a12a1.r1 Aspergillus nidulans 24hr asexual developmental and
 DEFINITION vegetative cDNA lambda zap library Emicella nidulans cdna clone
 n3a12a1 mRNA sequence.
 ACCESSION AA787396
 VERSION AA787396.1 GI:2847627
 KEYWORDS EST.
 SOURCE Emicella nidulans.
 ORGANISM Emicella nidulans.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiiales; Trichocomaceae; Emicella.
 1 (bases 1 to 541)
 Kuper, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
 Prade, R., and Roe, B.
 An Aspergillus nidulans EST Database
 Unpublished (1998)
 Other_ESTs: n3a12a1.f1
 Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
 Department of Chemistry and Biochemistry
 Advanced Center for Genome Technology, University of Oklahoma

RESULT 3
 AA783132 561 bp mRNA linear EST 29-JUL-1998
 LOCUS c11a1.r2 Aspergillus nidulans 24hr asexual developmental and
 DEFINITION vegetative cDNA lambda zap library Emicella nidulans cdna clone

QY	1120	ccggcggaagaacgaccccggaacgcggaagcaactcgtgctctcgtggctacccgacgaca	1179
Db	301	CAGCGACGACATGATGCCCTTATTAACCGGGGACCGGGTTGGTGTCCTCTACTCGACAGCT	360
QY	1180	gcagactgtcacagtaattcgtctccggaagcaactcggagcatatgcatctagcaag	1239
Db	361	CAGAGCTATACAAAGTTCACTCGGACCACTTACAAAGATCGGAATACTGGCCATTTCCAAG	420
QY	1240	atacaggaatcgtgacctaacgaagatgagccatctacaaga	1282
Db	421	ATTCCAGTTATCTCACTCTCCGGAATATCTCTTTTACAGCGA	463
RESULT	5		
LOCUS	AI211322	466 bp	mRNA linear EST 19-OCT-1998
DEFINITION	AI211322	06c07a1.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library	Emeritella nidulans cDNA clone 06c07a1 5', mRNA sequence.
ACCESSION	AI211322	GI:3773264	
VERSION	AI211322.1	GI:3773264	
KEYWORDS	EST.		
SOURCE	Emeritella nidulans.		
ORGANISM	Emeritella nidulans.		
REFERENCE	1 (bases 1 to 466)		
AUTHORS	Kuper,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,B.		
TITLE	An Aspergillus nidulans EST Database		
JOURNAL	Unpublished (1998)		
COMMENT	Other ESTs: 06c07a1.fl Contact: Bruce A. Roe, University of Oklahoma, broeou.edu Department of Chemistry and Biochemistry Advanced Center for Genome Technology, University of Oklahoma 620 Parrington Oval, Norman, OK 73019, USA Tel: 405 325 4912 Fax: 405 325 7762 Email: broeou.edu We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center Seq primer: T3 High quality sequence stop: 321. Location/Qualifiers 1..466 /organism="Emeritella nidulans" /strain="FGSC A26" /db_xref="taxon:162425" /clone="06c07a1" /clone_1lb="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library" /issue_type="vegetative mycelia asexual structures" /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"		
BASE COUNT	102 a	129 c	129 g 106 t
ORIGIN			
Query Match	10.9%;	Score 189.8;	DB 9; Length 466;
Best Local Similarity	70.4%;	Pred. No. 2.4e-19;	
Matches 254; Conservative	0;	Mismatches 107;	Indels 0; Gaps 0;
QY	129	tgagcagcagcatcattatctctcctcgtgacgacgcgattcgctcgacgagcgatcg	188
Db	104	TGGCCACCCAGTCGATTAATTTCTCTGACGAGCGATTCGTCGACGAGGATCG	163
QY	189	acgaetcgacttgaatcactcgagatcagaataactctgttggaacatgcagagc	248
Db	164	ACGACCGCTGCCCTTGATCTGGCTTCAACGAGATATCTCGCGGCGAGCTGGCAGGGCATC	223
QY	249	atcgacaagttggacatataccaggaatggtgcttcaagcagcatctgatacccccgt	308

QY	775	aagccgcaagcgtgactctatcgcgcgagcgtgctgcagcgtgatacgcgcctacactgtc	834
DB	224	ATTCACACAGCTGATTATATATCCAAAGACATAGGGATGATCCATCTGTGATTACACCTATC	283
QY	309	aacgagccagctgccccagacaccgcatatgtagatgcttacatgctcactgcgcgcg	368
DB	284	ACCGACAGATGTTCCCATGTCCACCGTGTGGACACGGGCTTACATGGCTACTGCGAGAAG	343
QY	369	gatactactctcgaagaaactcaggcagctcagatcagatctgaaggcgcctcttcg	428
DB	344	AACATATACGGTGTGCACACCAACACTGGGGACAGCCGACAGATATACAGGGCTGTGCGAG	403
QY	429	gcccttcagagaggggagatgtatctctatgctcgtatgctgtgttgtaacatatggcat	488
DB	404	GGCTCCATGATCGGGGCAATGATTCATGCTGGAGTGTTCGCAACCAATGTTCTTAT	463
QY	489	g	489
DB	464	G	464

QY	835	ccaccagaagcgtctcgaagcgcgctacgtgaactatcccatctacatccactctcaag	894
Db	64	CTTNNANNAATTATATGAGACGGGCGTCATGAACTACCCCATATATACCCCTTCGTGATG	123
QY	895	ccctcaagtcacactcgcgcgcagcagcagcagcctctacacatgcatcaacacgctcaat	954
Db	124	CGTTCAAGTCTCCGACGGGAGCATGTGGGATCTCTATACATNANGAGAGACGGGCT	183
QY	955	cgcagatgcacagactccacacatccctgggcacatcgtctgagaaccagacaacccaggt	1014
Db	184	CAAAATTGCGGATCTCACTACCTGCTGGAAACTTTATGAAANACCATGACATCTTCGAT	243
QY	1015	tcgctctctacaccaagacagacatagccctcgcgcagaacgctgcagcatctcatcctca	1074
Db	244	TCCCCAACTATACCTCCGATATGAAATCGGCCCAAAAAGTGCTCGTCTCTCTGTGA	303
QY	1075	acgacggaatcccatcatcactaagccgcgcgaagaacagcacta	1117
Db	304	CCGACGGAAATCCCTATTTATATGCGGGCCAGAGCAACCACTCA	346

	RESULT	7
AA786814		
LOCUS		
DEFINITION		
AA786814	376 bp	mRNA linear EST 31-JUL-1998
mc61l1a1.r1		Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emeritella nidulans CDNA clone mc61l1a1 5', mRNA sequence.

ACCESSION	AA786814
VERSION	AA786814.1
KEYWORDS	EST.
SOURCE	<i>Emmericella nidulans</i> .
ORGANISM	<i>Emmericella nidulans</i>

REFERENCE	TITLE
AUTHORS	JOURNAL
COMMENT	
Eukariotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emmentella. 1 (bases 1 to 376)	
Kuffer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,B.	An Aspergillus nidulans EST database unpublished (1998)
Contact: Bruce A. Roe, university of Oklahoma, broe@ou.edu	

Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: SK
High quality sequence stop: 367.

FEATURES	source	Location/Qualifiers
	1..376	
		/organism="Emmericella nidulans"
		/strain="PGSC A26"
		/db_xref="taxon:162425"
		/clone="m6c1a1"
		/clone_11b="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
		/tissue_type="vegetative mycelia, asexual structures"
		/note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT	92 a	98 c 91 g 95 t
ORIGIN		

[illegible]

Db	2	GACCATACAGCTCCTTGGACTGATCTGAAACCCACACACTCCGAGGTACCGCACTCTGG	61
Oy	675	taagactgggtggatcattgtatcgaactactccattgaagccctccgtatcgcagaca	734
Db	62	TATGCACTGGGTCAAGGACATCGTTCCCAATTACTGTGGGATGGGCTCCCAATTGGATTCC	121
Oy	735	gttaaacacg-tccagaagaactcttgcccggtgatacacaaacgcg-cagcgcttact	792
Db	122	GTCAAGCACCTTTGCAATATGATTTCTGGCCGGGTTATATTATGCCACACTGGGGCTACA	181
Oy	793	gtatcgcgaggtgctccgcagcgtgataccgccttaactgtgccctacagaaacgtaatg	852
Db	182	GCATCGSTGAAATCTACCATGGGGAACCGGAGTATACCTGGCCTTTACCAGAGTTATGTGG	241
Oy	853	acgcgcctactgaactatcccatltaetatacactccctcaacgccttaagctcaagctccg	912
Db	242	ACGGGGTCATGAACTATACCCATATATTACCCCTTCTGAATGGCTTCAAGTCCCTCGAGGC	301
Oy	913	gcagcatagcgcagccctcaacaatgatacaaacgcgttaaatcgcagctgttccagactcaa	972
Db	302	GGACCAATGTCGGATCTCTTAAACATGATCAACACAGTGCCTCAAAATTGTGGGATCTTA	361
Oy	973	caactccctggaca	984
Db	362	CACCTGCTTGGA	373

LOCUS	DEFINITION
AA785147	374 bp mRNA linear EST 28-JUL-1996
g5c1a1.r1	<i>Aspergillus nidulans</i> 24hr asexual developmental and
vegetative cDNA	<i>lambda zap</i> library <i>Emeticella nidulans</i> cDNA clone
g5c1a1.5'	mRNA sequence.

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AA/785147	AA/785147.1	GI:2845315	EST.	<i>Emerlicella nidulans</i> .	
				<i>Emerlicella nidulans</i> .	
				Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; <i>Emerlicella</i> .	
				(bases 1 to 374)	

AUTHORS Knappe, D., Gray, J., Hausner, J., Lai, H., Martin, W., Atamayo, R., Prade, R. and Roe, B.
TITLE An *Aspergillus nidulans* EST Database
JOURNAL Unpublished (1998)
COMMENT Other-ESTs: g5c10a1.f1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel.: 405 325 4812

Fax: 405 325 1762
 Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal
 Genetics Stock Center
 Seq primer: SK
 High quality sequence stop: 340.
 Location/Qualifiers
 1..374
 /organism="Emmericella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="g5c10a1"
 /clone_1lb="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zsp library"
 ...

BASE COUNT	XhoI: 5' end of cDNA cloned into EcoRI site of pluescript 3' end of cDNA cloned into XhoI site of pluescript"			
ORIGIN	84 a	107 c	99 g	83 t
				1 others

Db 108 TGGGCGACCGAGTCGATCTATTCCTCCTACGCGACGAGTTCCTGCGACGAGCGATG 167
 Qy 189 agactgcgactgtaactgcggtacagaaactgctggtggaacatgacgagcgc 248
 Db 168 AGGACCGCTGCTGTGATCTGCTCAACGAGATAGTCCGCGGCGAGCTGCGAGGCGATC 227
 Qy 249 atcgacaagtgtgactatccagggaatggcttcacacgcatctgcatcaccccggt 308
 Db 228 ATCAACACGAGTGTATATATCCAGACATGGGATTCACCTCCATCTGGATTACACCTATC 287
 Qy 309 acagccagcctgcgcacagacacgcacatgagatgcttaccatgctactgacgacgag 368
 Db 288 ACCGACGAGATTCCTCCGATGTCACCGCTGTTGGAAACGGGCTTCATGCTACTGGCAAG 347
 Qy 369 gataatac 377
 Db 348 AACATATAC 356

RESULT 11
 A1212282 389 bp mRNA linear EST 19-OCT-1998
 LOCUS x1f041.r1 Aspergillus nidulans 24hr asexual developmental and
 DEFINITION vegetative CDNA lambda zap library Emericella nidulans cDNA clone
 x1f041 5', mRNA sequence.

ACCESSION A1212282
 VERSION A1212282.1 GI:3774224
 KEYWORDS EST.
 SOURCE Emericella nidulans.
 ORGANISM Emericella nidulans.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiiales; Trichocomaceae; Emericella.

REFERENCE 1 (bases 1 to 389)
 AUTHORS Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R.,
 Prade,R. and Roe,B.
 TITLE An Aspergillus nidulans EST Database
 JOURNAL Unpublished (1998)
 COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
 Department of Chemistry and Biochemistry
 Advanced Center for Genome Technology, University of Oklahoma
 620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal
 Genetics Stock Center
 Seq primer: T3
 High quality sequence stop: 307.

FEATURES

source

1..389
 Location/Qualifiers
 /organism="Emericella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="x1f041"
 /clone.lib="Aspergillus nidulans 24hr asexual
 developmental and vegetative CDNA lambda zap library"
 /tissue.type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
 3' end of cDNA cloned into XhoI site of pBluescript"
 BASE COUNT 95 a 116 c 82 g 96 t
 ORIGIN

Query Match 7.7%; Score 133; DB 9; Length 389;
 Best Local Similarity 58.9%; Pred. No. 8.5e-11;
 Matches 229; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

Qy 1062 ttatcatctcaacagcgaatcccatcatctacgcgcgcaagaacagcactagcc 1121
 Db 1 TTCTCTTTTGTGACCGAGAACTCCTATTGCTTATGCGGCGCAGAGACACACTATTC 60

Qy 1122 ggcggaacagccccgcgcgaacgcgaacactgctctgggtactaccgacgcagc 1181
 Db 61 GGCAGAAATGATCCCTATTAACCGGAGCGGTTGTGGTCTCTACTCAGCAGCTCA 120
 Qy 1182 gactcgtaaaatattatgctccgcggaacgcatccggaactagccattagcaaat 1241
 Db 121 GAGCTATACAAATTCATCGCGACCACTAACAGATCCGAAACTGGCCATTTCCAAAT 180
 Qy 1242 acaggaatcgtagcctacacgaactgcgcacatctacaagaacgacaacatgcgcat 1301
 Db 181 TCCAGTTATCTCACTTCCCGGAATACTCCTTTTACAGGATAGCACTATATCGCATG 240
 Qy 1302 cgcgaagggcagatggtgcgagatcgtagctatctgtccacaaggggtctcggt 1361
 Db 241 CGCAAGGCTCTGGGGGCTCTCAGGCTCTCACTCTCTCAACAAATATCGGCAACGATC 300
 Qy 1362 gattcgtaacctctcctctggtggtggttcacacgcgcgcgaacatgacgag 1421
 Db 301 GGTTCCTATACATTCGACCTGATGATCAATGATACACAGCGGCGCCCACTAGTGGA 360
 Qy 1422 gtcatcgctgcagacgcgtgaaggttg 1450
 Db 361 CTGTACACATGCTCTCTGTCCAGGTCG 389

RESULT 12
 AA784386 358 bp mRNA linear EST 29-JUL-1998
 LOCUS d5b1a1.r1 Aspergillus nidulans 24hr asexual developmental and
 DEFINITION vegetative CDNA lambda zap library Emericella nidulans cDNA clone
 d5b1a1 5', mRNA sequence.

ACCESSION AA784386
 VERSION AA784386.1 GI:2844554
 KEYWORDS EST.
 SOURCE Emericella nidulans.
 ORGANISM Emericella nidulans.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiiales; Trichocomaceae; Emericella.

REFERENCE 1 (bases 1 to 358)
 AUTHORS Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R.,
 Prade,R. and Roe,B.
 TITLE An Aspergillus nidulans EST Database
 JOURNAL Unpublished (1998)
 COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
 Department of Chemistry and Biochemistry
 Advanced Center for Genome Technology, University of Oklahoma
 620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal
 Genetics Stock Center
 Seq primer: SK
 High quality sequence stop: 348.

FEATURES

source

1..358
 Location/Qualifiers
 /organism="Emericella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="d5b1a1"
 /clone.lib="Aspergillus nidulans 24hr asexual
 developmental and vegetative CDNA lambda zap library"
 /tissue.type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
 3' end of cDNA cloned into XhoI site of pBluescript"
 BASE COUNT 79 a 104 c 97 g 78 t
 ORIGIN

Query Match 7.6%; Score 131.8; DB 9; Length 358;
 Best Local Similarity 72.0%; Pred. No. 1.3e-10;
 Matches 172; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY	129	tggcgcgcgaacccatttatttccttccttcacgagatcgatttgcaaggacgatgtg	188
Db	120	TGGCGCAGCCGATGCATCTATTTCCTTCCGTACGCCAGCACATGCTGCACGGACGATG	179
QY	189	acgactgtagctgttaatactcgtgcgaatcagaatactcgtgtgtgaacaatgycagggcatc	248
Db	180	ACGACCGCTCCTGTGATCGCTCAACGAGAGATCATCGTGCGCGCGACACTGGCAGGCAATC	239
QY	249	atcgcaaatgttgatcataataccaaggaaatggygcttcacagcatctgtacacccccgt	308
Db	240	ATCAACCACGTTGATTAATATCCAAAGACAATGGGATTCACATCGCATCTGATTAACACTATC	299
QY	309	accagcccacgtgcccccagccagccagatatgagtgccttacatgagctaacgtgcga	367
Db	300	ACCGACGACATTCCTCCGATGTCAACGCTGTTGGAGACGGGCTTCACATGCTAATGCGAGAA	358
RESULT	13		
AI210530			
LOCUS	17h08a1.r1	Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone 17h08a1 5', mRNA sequence.	
DEFINITION	AI210530		
ACCESSION	AI210530		
VERSION	AI210530.1		
KEYWORDS	GI:3772472		
SOURCE	EST.		
ORGANISM	Emericella nidulans.		
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eutroliales; Trichocomaceae; Emericella.		
AUTHORS	1 (bases 1 to 304) Kuper,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,B. An Aspergillus nidulans EST Database Unpublished (1998)		
TITLE	Other_ESTS: 17h08a1.fl		
JOURNAL	Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu		
COMMENT	Department of Chemistry and Biochemistry Advanced Center for Genome Technology, University of Oklahoma 620 Parrington Oval, Norman, OK 73019, USA Tel: 405 325 4912 Fax: 405 325 7762 Email: broe@ou.edu We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center Seq primer: T3 High quality sequence stop: 293. Location/Qualifiers 1..304 /organism="Emericella nidulans" /strain="FGSC A26" /db_xref="taxon:162425" /clone="17h08a1" /clone.lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library" /issue.type="vegetative mycelia, asexual structures" /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"		
BASE COUNT	79 a	89 c	63 g
ORIGIN	73 t		
Query Match	7.6%	Score 131.2;	DB 9; Length 304;
Best Local Similarity	64.5%;	Pred. No. 1.7e-10;	
Matches 196;	Conservative 0;	Mismatches 108;	Indels 0; Gaps 0;
QY	945	accgtcaaatccgcatgtctcagactcaacaactctgtgcacatctgcgaagaaccagac	1004
Db	1	ACAGTCGGCTCAAAATGTCGGGATCTCAACACTGCTTGGAAMACTTATTCGACAAACATGAC	60
QY	1005	aaccaacggttcgctttcttaccacacagacatagacctcgccaagaacgtgcagcatc	1064

QY	1065	atcatctcaagacgggaatccccatcatcttaagccggcgcaagaagaacgactacgcgcgc	1124
Db	61	AAATCTCGATTTCCCAACTATTAATCCCGGATATAGTGGCGCAAGAACGCTCTCGCTTC	120
QY	1125	ggaacagaccccggaacgcgaagaacacttgcctcctcgcgcgtacccgacgacgacg	1184
Db	181	AGCAATGATCCCTATAACCGGGAGCCGGTGGTGAGTCTCTACACGACACACTCAGAG	240
QY	1185	ctgtacaagttaattgtcctcgcgcgaacgcaatccgcgaactaactcattacgaagata	1244
Db	241	CTATACAAAGTTTCATCGCGACCTACTAACAGATCCGAAACTGGCATTTCCTCAAAAGATTCC	300
QY	1245	ggat 1248	
Db	301	AGTT 304	
RESULT	14		
LOCUS	AA787924	405 bp	mRNA linear EST 31-JUL-1996
DEFINITION	r2h03ai.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone r2h03ai 5', mRNA sequence.		
ACCESSION	AA787924		
VERSION	AA787924.1	GI:2848155	
KEYWORDS	EST.		
SOURCE	Emericella nidulans.		
ORGANISM	Emericella nidulans.		
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.		
AUTHORS	1 (bases 1 to 405) Kupfer,D., Gray,J., Hausener,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,J.B.		
TITLE	An Aspergillus nidulans EST Database		
JOURNAL	Unpublished (1998)		
COMMENT	Other ESTs: r2h03ai.fl Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu Department of Chemistry and Biochemistry Advanced Center for Genome Technology, University of Oklahoma 620 Parrington Oval, Norman, OK 73019, USA Tel: 405 325 4912 Fax: 405 325 7762 Email: broe@ou.edu We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center Seq primer: SK High quality sequence stop: 321. Location/Qualifiers 1. 405 /organism="Emericella nidulans" /strain="FGSC A26" /db_xref="taxon:162425" /clone="r2h03ai" /clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library" /tissue_type="vegetative mycelia, asexual structures" /note="Vector: pluescript SK-; Site: 1: EcoRI; Site: 2: XhoI; 5' end of cDNA cloned into EcoRI site of pluescript 3' end of cDNA cloned into XhoI site of pluescript"		
BASE COUNT	99 a 124 c 81 g 94 t		7 others
ORIGIN			
Query Match	6.7%:	Score 116.2;	DB 9; Length 405;
Best Local Similarity	55.9%:	Pred. No.2.8e-08;	
Matches	214;	Conservative 0;	Mismatches 169; Indels 0; Gaps 0;
QY	1077	gacgaatcccatcatctacgcgcgcgaagaacgacactcgcgcgcgaacgaccc	1136
Db	19	GCCGATCCCATTTATNGTATNCCGGGCAAGANCAACACTTTCACCGGANCAAGATCCG	78

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 13:44:19 ; Search time 293.75 Seconds
(without alignments)
10134.909 Million cell updates/sec

Title: US-09-710-339-1

Perfect score: 1734
1 tccatcaagctctcccttc.....aaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1734	100.0	1734	22	AAF90208
2	1707	98.4	1914	21	AAF12832
3	1063.8	61.3	2337	11	AAQ06786
4	806	46.5	1389	21	AAF11252
5	570	32.9	1956	16	AAQ88712
6	439	25.3	1389	21	AAF11252
7	378.6	21.8	1404	13	AAQ25079
8	377.4	21.8	4214	8	AAQ70916
9	375.4	21.6	1404	15	AAQ7668

10	375.4	21.6	1404	15	AAQ77665
11	374.2	21.6	483	21	AAF11314
12	373.8	21.6	1404	15	AAQ77667
13	373.8	21.6	1404	15	AAQ77666
14	369.2	21.3	2289	11	AAQ06388
15	368.8	21.3	605	21	AAF11273
16	349.6	20.2	2290	9	AAQ81525
17	348	20.1	2290	9	AAQ81477
18	344.2	19.9	1048	21	AAF11253
19	336.4	13.6	619	21	AAF12538
20	220.6	12.7	292	21	AAF11355
21	206.6	11.9	1602	21	AAF07664
22	190	11.0	1067	10	AAQ91274
23	190	11.0	1316	8	AAQ70507
24	190	11.0	1316	13	AAQ24569
25	190	11.0	1316	22	AAQ59332
26	190	11.0	1596	10	AAQ91273
27	178	10.3	1315	10	AAQ91074
28	165.4	9.5	592	21	AAF11268
29	143	8.2	3311	17	AAQ10562
30	143	8.2	3311	19	AAQ61458
31	143	8.2	7432	19	AAQ61459
32	129.2	7.5	173	21	AAF11472
33	92	5.3	1596	10	AAQ91273
34	70.6	4.1	839	21	AAF11277
35	68.6	4.0	93	19	AAQ15070
36	68.4	3.9	922	19	AAQ15072
37	68.4	3.9	922	19	AAQ15073
38	68.4	3.9	925	19	AAQ15076
39	68.4	3.9	928	19	AAQ15074
40	68.4	3.9	1438	10	AAQ91275
41	65.2	3.8	256	17	AAQ10566
42	63.4	3.7	915	19	AAQ15075
43	62.4	3.6	7417	23	AAQ16368
44	59.4	3.4	1067	10	AAQ91274
45	58	3.3	2582	18	AAQ73117

ALIGNMENTS

RESULT	1
ID	AAF90208
XX	AAF90208 standard; DNA; 1734 BP.
AC	AAF90208;
XX	
DT	06-AUG-2001 (first entry)
XX	
DE	Nucleotide sequence of a fungamyl-1-like alpha-amylose.
XX	
KW	Fungamyl-1-like alpha-amylose; glucanase; dextrinase; maltose;
KW	alcohol; starch; dough improver; brewing; starch liquification; ss.
XX	
OS	Aspergillus oryzae.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "fungamyl-1-like alpha-amylose"
FT	54..113
FT	/*tag= b
FT	114..1547
FT	mat_peptide
FT	/*tag= c
XX	
PN	WO200134784-A1.
XX	
PD	17-MAY-2001.
XX	
PF	10-NOV-2000; 2000MO-DK00626.
XX	
PR	10-NOV-1999; 99DK-0001617.
XX	

Variant alpha amy1
Aspergillus niger
Variant alpha amy1
Variant alpha amy1
AM1 gene encoding
Aspergillus niger
Sequence encoding
5' end of alpha-am
Aspergillus niger
Aspergillus oryzae
Aspergillus niger
Fusarium venenatum
Plasid NA2 DNA en
Sequence of the TA
Sequence of TAKA-a
Aspergillus oryzae
Plasid NA1 DNA en
Sequence of Asperg
Aspergillus niger
Alpha-amylose gene
Alpha-amylose amyA
Vector pPR70 conta
Aspergillus niger
Plasid NA1 DNA en
Aspergillus niger
Humicola insolens
Hybrid DNA compri
Hybrid DNA compri
Hybrid DNA compri
DNA encoding Asper
Alpha-amylose gene
Hybrid DNA compri
Diosphila melanog
Plasid NA2 DNA en
Actinoplanes sp. a

PA (NOVO) NOVOZYMES AS.
XX
XX Bisgard-Frantzen H, Svendsen A, Pedersen S;
XX WPI: 2001-367478/38.
DR P-PSDB: AAB84206.
XX
XX New variant of Fungamyl-1-like alpha-amylase, useful for production of
PT maltose syrups, includes mutations that improve stability against heat
PT and acidic pH -
XX
XX Example 1; Page 40-42; 49pp; English.
PS
XX
XX The present sequence encodes a fungamyl-1-like alpha-amylase. The
CC specification describes variants of this fungamyl-1-like alpha-amylase,
CC which have an alteration in one the amino acid regions 98-110, 150-160,
CC 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion
CC or substitution of an amino acid or an insertion of an amino acid
CC downstream of a particular position. The variants retain alpha-amylase
CC activity, and have better heat stability and/or stability at acidic pH,
CC relative to wild-type enzyme. The variants can therefore be used at
CC higher temperatures (more efficient conversion or faster reaction, and
CC have reduced need for cooling and reduced risk of contamination). The
CC variants may also be used in conjunction with other enzymes,
CC particularly glucoamylase during dextrinisation. The variants are
CC used to produce syrups, particularly of high maltose content, or alcohol,
CC from starch; as dough improver for baked goods; in brewing, to increase
CC fermentability of the wort; and for liquefaction of starch.
XX
XX Sequence 1734 BP; 450 A; 460 C; 425 G; 399 T; 0 other:
SQ

Query Match 100.0%; Score 1734; DB 22; Length 1734;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tcaatcaagaactcctcctctctgtgaacaataaaccacagaaggcatattatgtgtcgcg 60
DB 1 tcaatcaagaactcctcctctctgtgaacaataaaccacagaaggcatattatgtgtcgcg 60
QY 61 cgtgtgtgtctatttctgttagcgttcagtgcgcgacgtccttgggtgtgaagc 120
DB 61 cgtgtgtgtctatttctgttagcgttcagtgcgcgacgtccttgggtgtgaagc 120
QY 121 cgtgtgtgtgtgacgtgcacatcatttatttctctctcagcgatgattgcaagagcg 180
DB 121 cgtgtgtgtgtgacgtgcacatcatttatttctctctcagcgatgattgcaagagcg 180
QY 181 atgggtcgaagactgacgtgttaactgtcggtatcagaataactgtgtgtgaacatgagc 240
DB 181 atgggtcgaagactgacgtgttaactgtcggtatcagaataactgtgtgtgaacatgagc 240
QY 241 aagggtcatcgacaagtgtgactatattcagaagaaatgggtcttcacgcacatgtgatca 300
DB 241 aagggtcatcgacaagtgtgactatattcagaagaaatgggtcttcacgcacatgtgatca 300
QY 301 cccccgttcaacgacgtgcgccccagaccacgcgcatatggagatgctccacatgtgctact 360
DB 301 cccccgttcaacgacgtgcgccccagaccacgcgcatatggagatgctccacatgtgctact 360
QY 361 ggcagcagagatatatactctctgaacgaaaactacgcgacatgcagatgacttgaagcgcg 420
DB 361 ggcagcagagatatatactctctgaacgaaaactacgcgacatgcagatgacttgaagcgcg 420
QY 421 tctcttcggcccttcagtgaaggggagatgtatcttatgttcgaatgtgtgtctaacacata 480
DB 421 tctcttcggcccttcagtgaaggggagatgtatcttatgttcgaatgtgtgtctaacacata 480
QY 481 tgggtcatatagtagcggttagctcagtgatctacagtgatgtgtttaacgcgttcagtcc 540
DB 481 tgggtcatatagtagcggttagctcagtgatctacagtgatgtgtgtttaacgcgttcagtcc 540
QY 541 aagactactcaccacgcttcgttcatcacaactatgaatcagatcagctcaggttgagc 600

DB 541 aagactactcaccacgcttcgttcatcacaactatgaatcagatcagctcaggttgagc 600
QY 601 attgtgtcctagagagataaacaactgtctctctgtcctatctcgtatccaccagaagatgtg 660
DB 601 attgtgtcctagagagataaacaactgtctctctgtcctatctcgtatccaccagaagatgtg 660
QY 661 tcaagaatgatgtgtagcagctgtggtgagatccttggtatcgtatcgaactactccatgtagcgcg 720
DB 661 tcaagaatgatgtgtagcagctgtggtgagatccttggtatcgtatcgaactactccatgtagcgcg 720
QY 721 tccgtatcgacaacaglaaaacacgltccagaagaagacttgcgtccgggttaacaaagcgcg 780
DB 721 tccgtatcgacaacaglaaaacacgltccagaagaagacttgcgtccgggttaacaaagcgcg 780
QY 781 cagcgctgtactgtatctgcgagagtgctgcagcggtgataccggtactacttgcctacc 840
DB 781 cagcgctgtactgtatctgcgagagtgctgcagcggtgataccggtactacttgcctacc 840
QY 841 agaacgctatggagcgggtactgaactatccattactactccactccctcaagccttca 900
DB 841 agaacgctatggagcgggtactgaactatccattactactccactccctcaagccttca 900
QY 901 agtcaacctccggaagcatgtgagcgaacctatacaatgatcaacacgcgttcaaatccgact 960
DB 901 agtcaacctccggaagcatgtgagcgaacctatacaatgatcaacacgcgttcaaatccgact 960
QY 961 gtccaagcttaaaactctgtgagacatctgttgagaacccagcaacccacggttgcgt 1020
DB 961 gtccaagcttaaaactctgtgagacatctgttgagaacccagcaacccacggttgcgt 1020
QY 1021 cttaacccaacgacatagccctcgcgaagaacgttcgaagatcatcatcctcctcaagcagc 1080
DB 1021 cttaacccaacgacatagccctcgcgaagaacgttcgaagatcatcatcctcctcaagcagc 1080
QY 1081 gaatcccatcatctacgcgcgcgaagaacagcactacgcgcgcgaagaacgcgcgcga 1140
DB 1081 gaatcccatcatctacgcgcgcgaagaacagcactacgcgcgcgaagaacgcgcgcga 1140
QY 1141 accgcaagaagcaactgtgtctgtgcgtaccgcgcgcgaagaagcgtgttaagaattatgt 1200
DB 1141 accgcaagaagcaactgtgtctgtgcgtaccgcgcgcgaagaagcgtgttaagaattatgt 1200
QY 1201 ccttcgcgaacgcacatccggaactatgtccaatagcaagaataaggaattcgttgacctata 1260
DB 1201 ccttcgcgaacgcacatccggaactatgtccaatagcaagaataaggaattcgttgacctata 1260
QY 1261 aagaactgtgccatctacaagaacgacaacagatcgcacatgccaagggcagacagatgggt 1320
DB 1261 aagaactgtgccatctacaagaacgacaacagatcgcacatgccaagggcagacagatgggt 1320
QY 1321 cgcagatcgtgactatcttccaagaaggtgtcctcgggttgatctgatatccctctct 1380
DB 1321 cgcagatcgtgactatcttccaagaaggtgtcctcgggttgatctgatatccctctct 1380
QY 1381 tgaatgtgtcgggtttacacagccgcgcgaagaatgtgacggagtcattgtgttcagcagcg 1440
DB 1381 tgaatgtgtcgggtttacacagccgcgcgaagaatgtgacggagtcattgtgttcagcagcg 1440
QY 1441 tgaacggttgcgtcggatgtgaatgtgcctgttccatgtgcaggttgggtctacatggtat 1500
DB 1441 tgaacggttgcgtcggatgtgaatgtgcctgttccatgtgcaggttgggtctacatggtat 1500
QY 1501 tgtatccgactcggaaagtgtgcaggttagcagaagatcgttaatcgttgaaggttgagaga 1560
DB 1501 tgtatccgactcggaaagtgtgcaggttagcagaagatcgttaatcgttgaaggttgagaga 1560
QY 1561 gtatatgatgtactgtatcatcaactcgtgcattggaaaggttgaggtttgagttgata 1620
DB 1561 gtatatgatgtactgtatcatcaactcgtgcattggaaaggttgaggtttgagttgata 1620
QY 1621 gttggaatcgttactgtcgtcatcccttataactctcgaattgttttcgaaacctaagt 1680
DB 1621 gttggaatcgttactgtcgtcatcccttataactctcgaattgttttcgaaacctaagt 1680

Db 1621 gttggatcgtcttcgtcgtcgtccatccctattcctcttgatgttttcgaacctaatg 1680

Qy 1681 ccaagccgcgtagtctattataggaataaaaaaaaaaaaaaaaaaaaaa 1734
|||||
Db 1681 ccaagcagcagtctattataggaataaaaaaaaaaaaaaaaaaaaaa 1734

RESULT 2

AAFI2832

ID AAFI2832 standard; cDNA: 1914 BP.

AAFI2832:

13-MAR-2001 (first entry)

Aspergillus oryzae EST SEQ ID NO:5355.

Multiple gene expression; filamentous fungal cell; EST:
expressed sequence tag; Fusarium venenatum; Aspergillus niger;
Aspergillus oryzae; Trichoderma reesei; identification; recombination;
culture condition; environmental stress; spore morphogenesis;
metabolic pathway engineering; catabolic pathway engineering; ss.
Aspergillus oryzae.

WO200056762-A2.

28-SEP-2000.

22-MAR-2000; 2000MO-US07781.

22-MAR-1999; 99US-0273623.

(NOVO) NOVO NORDISK BIOTECH INC.
(NOVO) NOVO NORDISK AS.

Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
WPI: 2000-594572/56.

Monitoring differential expression of genes in filamentous fungal cells
uses fluorescence-labeled nucleic acids isolated from the cells and a
substrate of expressed sequence tags -

Claim 88; Page 2231-2232; 3161pp; English.

The present invention describes a method for monitoring differential
expression of genes in a first filamentous fungal (FF) cell relative to
expression of the same genes in one or more second filamentous fungal
cells. The method uses fluorescence-labeled nucleic acids isolated from
the FF cells and a substrate of expressed sequence tags (EST). The ESTs
are used in the methods for monitoring differential expression of genes
in a first filamentous fungal (FF) cell relative to expression of the
same genes in one or more second filamentous fungal cells. Monitoring
the global expression of genes from FF cells allows the production
potential of the microorganisms to be improved. New genes may be
discovered, possible functions of unknown open reading frames can be
identified and gene copy number variation and stability can be
monitored. The expression of genes can be used to study how FF cells
adapt to changes in culture conditions, environmental stress, spore
morphogenesis, recombination, metabolic or catabolic pathway
engineering. Using ESTs provides several advantages over genomic or
random cDNA clones including elimination of redundancy as one spot on an
array equals one gene or open reading frame, and organisation of the
microarrays based on function of the gene products to facilitate
analysis of the results. AAFI248 to AAFI1247 represents ESTs from
Fusarium venenatum; AAFI1248 to AAFI1853 represents ESTs from Aspergillus
niger; AAFI1854 to AAFI4878 represents ESTs from Aspergillus oryzae; and
AAFI4879 to AAFI5337 represents ESTs from Trichoderma reesei, which are
all specifically claimed in the present invention.

Sequence 1914 BP; 484 A; 510 C; 461 G; 459 T; 0 other;

Query Match	98.4%;	Score 1707;	DB 21;	Length 1914;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 1707;	Conservative	0;	Mismatches	0;	Indels
					Gaps
					0;
QY	1	tcaacataaactctctctctctctctgaacaataaaccacagaagcatatgatgtgcg	60		
Db	3	tcaacataaactctctctctctctctgaacaataaaccacagaagcatatgatgtgcg	62		
QY	61	ctgtgtgtctctattctctgtacggccttcaggtctcgcgacctgtgtgtgtcgaacgc	120		
Db	63	ctgtgtgtctctattctctgtacggccttcaggtctcgcgacctgtgtgtgtcgaacgc	122		
QY	121	ctcgcgacgtgcgcatcgcaatccattttctctctctacagcatcggtttcgcaagacgcg	180		
Db	123	ctcgcgacgtgcgcatcgcaatccattttctctctctacagcatcggtttcgcaagacgcg	182		
QY	181	atggtctgaagacatcgagactgtlaaactcgtgcagatcaaaataactgtgtgtgaacatgvcg	240		
Db	183	atggtctgaagacatcgagactgtlaaactcgtgcagatcaaaataactgtgtgtgaacatgvcg	242		
QY	241	agggcatactcgaacaaagtgtgactataccaggaatgvgcttccacagccatctgtatca	300		
Db	243	agggcatactcgaacaaagtgtgactataccaggaatgvgcttccacagccatctgtatca	302		
QY	301	cccccggttaacggccagcgtgcggcccgacacccggacatagtgagatgacctaccatgvcgact	360		
Db	303	cccccggttaacggccagcgtgcggcccgacacccggacatagtgagatgacctaccatgvcgact	362		
QY	361	ggcgcgcagataatactctctctgaagaaataacgcgcacgcacgcagatgttgaaagcgc	420		
Db	363	ggcgcgcagataatactctctctgaagaaataacgcgcacgcacgcagatgttgaaagcgc	422		
QY	421	tctcttcgcgccttcataatgaaaggggagatgatacttaatgtgtcgaatgtgtgtgtaacata	480		
Db	423	tctcttcgcgccttcataatgaaaggggagatgatacttaatgtgtgtcgaatgtgtgtgtaacata	482		
QY	481	tggcgataatgtgaagcggtgaagctcaagatctgaagtggttttaaacgcgttcaagttccc	540		
Db	483	tggcgataatgtgaagcggtgaagctcaagatctgaagtggttttaaacgcgttcaagttccc	542		
QY	541	aagactactccacccggttctgtttcaatcaaaactatgaagaatcgaatcgaattgaaag	600		
Db	543	aagactactccacccggttctgtttcaatcaaaactatgaagaatcgaatcgaattgaaag	602		
QY	601	attctgtgcctcagaagataaacaactgtctctcttcgcgcgactctcgatatacccaagaagtgtg	660		
Db	603	attctgtgcctcagaagataaacaactgtctctcttcgcgcgactctcgatatacccaagaagtgtg	662		
QY	661	tcaagaatgtgaatgtgaagactgtgtgtgtatctgtatcgaactatcccatctgaacgcgc	720		
Db	663	tcaagaatgtgaatgtgaagactgtgtgtgtatctgtatcgaactatcccatctgaacgcgc	722		
QY	721	tccglatatcgacaacagtaaaacacgctccagaagactctlgcccggtgtatacaaaacgcg	780		
Db	723	tccglatatcgacaacagtaaaacacgctccagaagactctlgcccggtgtatacaaaacgcg	782		
QY	781	caggcggtgtactgtatgt	840		
Db	783	caggcggtgtactgtatgt	842		
QY	841	agaagctatagtgaacggcgctactgaactatcccatcttactatccatctcccaacgccttca	900		
Db	843	agaagctatagtgaacggcgctactgaactatcccatcttactatccatctcccaacgccttca	902		
QY	901	agttcaacctccggcagcatgtgacgcacctataacaatgtatcaaacgcgttcaaatccgact	960		
Db	903	agttcaacctccggcagcatgtgacgcacctataacaatgtatcaaacgcgttcaaatccgact	962		
QY	961	gttccagactcaaacctctctgtgcacatctgttctgagaacacagaacacccacgcgttgcctt	1020		
Db	963	gttccagactcaaacctctctgtgcacatctgttctgagaacacagaacacccacgcgttgcctt	1022		

```
OY 1021 cttaaccaagacataagcctctgcgaagaagctgcagcatctcatctcaacgacg 1080
      |||||||
DB 1023 cttaaccaagacataagcctctgcgaagaagctgcagcatctcatctcaacgacg 1082
OY 1081 gaatcccatcatctagcgcgcgcgaagaacagcactacgcgcgcgcgaaccgcgcga 1140
      |||||||
DB 1083 gaatcccatcatctagcgcgcgcgaagaacagcactacgcgcgcgcgaaccgcgcga 1142
OY 1141 accgcgaagcaacctggtcttgcgtctccgcacgcagagagagctgtacaaagttaattg 1200
      |||||||
DB 1143 accgcgaagcaacctggtcttgcgtctccgcacgcagagagctgtacaaagttaattg 1202
OY 1201 cctccgcgaacgcaatccgcgaactatgccaattagcaagaatatacagagattcgtgacctaca 1260
      |||||||
DB 1203 cctccgcgaacgcaatccgcgaactatgccaattagcaagaatatacagagattcgtgacctaca 1262
OY 1261 agaatgtgcccatactacaagaagacacacacagatgcacatgcgcgaaggacagatgggt 1320
      |||||||
DB 1263 agaatgtgcccatactacaagaagacacacacagatgcacatgcgcgaaggacagatgggt 1322
OY 1321 cgcgaatcgtgactatcttctccaaagggtgctggtggtgattcgtatccctctcct 1380
      |||||||
DB 1323 cgcgaatcgtgactatcttctccaaagggtgctggtggtgattcgtatccctctcct 1382
OY 1381 tgaagtgtgcgggttacacacagccgcgcagcaattgacgcgaaggtcatctgctgcacgacgc 1440
      |||||||
DB 1383 tgaagtgtgcgggttacacacagccgcgcagcaattgacgcgaaggtcatctgctgcacgacgc 1442
OY 1441 tgaagtggtgttcggaatggaatgtgtcctgttctctatgycagagtggtgctacctaagggtat 1500
      |||||||
DB 1443 tgaagtggtgttcggaatggaatgtgtcctgttctctatgycagagtggtgctacctaagggtat 1502
OY 1501 tctaccgactgagaagttgacggtacgacgaatctgtagtagctgcgcgaagggtgaga 1560
      |||||||
DB 1503 tctaccgactgagaagttgacggtacgacgaatctgtagtagctgcgcgaagggtgaga 1562
OY 1561 gtatatgtagtactgtctatccaatctgycatltgycacgtgagttgagttgatlga 1620
      |||||||
DB 1563 gtatatgtagtactgtctatccaatctgycatltgycacgtgagttgagttgatlga 1622
OY 1621 gttgagtcgttactgctgtcatcccttatactcttgcattgttttcgaaacctaatg 1680
      |||||||
DB 1623 gttgagtcgttactgctgtcatcccttatactcttgcattgttttcgaaacctaatg 1682
OY 1681 ccaagcagctagctatcatatagagaa 1707
      |||||||
DB 1683 ccaagcagctagctatcatatagagaa 1709

RESULT 3
AA006786
ID AA006786 standard; DNA; 2337 BP.
XX
AC AA006786;
XX
DT 06-MAR-1991 (first entry)
XX
DE Taka-amy1ase A gene.
XX
KW Alcohol; ss.
XX
OS Aspergillus oryzae RIB 40.
XX
PN JP02268685-A.
XX
PD 02-NOV-1990.
XX
PF 07-APR-1989; 89JP-0086787.
XX
PR 07-APR-1989; 89JP-0086787.
XX
PA (JOZO-) JOZO SHIGEN KENKYUS.
XX
```

```
DR WPI; 1990-372009/50.
XX
PT New gene, vector and transformant of taka amy1ase A - isolated
PT from Aspergillus oryzae, used to produce alcohol.
XX
PS Claim 1; Fig 1; 6pp; Japanese.
XX
CC The DNA was sequenced from a clone isolated from a A. oryzae RIB 40
CC genomic library. It can be used to express recombinant
CC taka-amy1ase useful in the prodn. of alcoholic drinks and alcohols.
CC See also AA006784 and AA006785.
XX
SQ Sequence 2337 BP; 630 A; 600 C; 533 G; 574 T; 0 other;

Query Match 61.3%; Score 1063.8; DB 11; Length 2337;
Best Local Similarity 75.3%; Pred. No. 1.8e-237;
Matches 1700; Conservative 0; Mismatches 7; Indels 552; Gaps 8;

OY 1 tcaatcaagctctccctctctgaacaataaaccacagaagagattatgattgctg 60
      |||||||
DB 68 tcaatcaaacctctccctctctgaacaataaaccacagaagagattatgattgctg 127
OY 61 cgtgtgtgtctctatctctgtacgcctcgaagtcgcgcgcacgtgttgcgtgcaacgc 120
      |||||||
DB 128 cgtgtgtgtctctatctctgtacgcctcgaagtcgcgcgcacgtgttgcgtgcaacgc 187
OY 121 ctgcgcagctgcgcagatgcgaatccattattctctctcaagatcgatttgcagaagcg 180
      |||||||
DB 188 ctgcgcagctgcgcagatgcgaatccattattctctctcaagatcgatttgcagaagcg 247
OY 181 atggttcgacgactgcgcactgtgaatcctgcgcgac----- 216
      |||||||
DB 248 atggttcgacgactgcgcactgtgaatcctgcgcgac----- 307
OY 217 -----agaaatagtgtgtgaaacatgycagggc 245
      |||||||
DB 308 tcagaagaagagatgtgaactgactgtgatagaactactggtggaacatgcgcagggc 367
OY 246 atcatcgaca----- 255
      |||||||
DB 368 atcatcgacaaggtlaaatgccccttatacaaaaaaagaagaagcaagaagaaaaa 427
OY 256 -----agttggaatataccagggatgg 280
      |||||||
DB 428 taataataaagaacctgtgcttaacctacacatggttgatataccaaggaatgg 487
OY 281 cttaacagccatctgcatcaccccgcttacagccagctgcgcgcagacacacgcatatgg 340
      |||||||
DB 488 cttaacagccatctgcatcaccccgcttacagccagctgcgcgcagacacacgcatatgg 547
OY 341 agatgctcaccatgctactgycagcagagatat----- 373
      |||||||
DB 548 agatgctcaccatgctactgycagcagagatat----- 607
OY 374 -----atacctctg 383
      |||||||
DB 608 tctactgtcatctatttctacatcatatgaactaactgtatggttttagaactctctg 667
OY 384 aacgaaaactacgacatgcagatgagatgaaaggcgtctctctgcgccttcaatgagag 443
      |||||||
DB 668 aacgaaaactacgacatgcagatgagatgaaaggcgtctctctgcgccttcaatgagag 727
OY 444 gggatgtatcttatggtgcagtgtgtgtgttaaccatat----- 481
      |||||||
DB 728 gggatgtatcttatggtgcagtgtgtgtgttaaccatat----- 787
OY 482 -----gggatagtatgg 495
      |||||||
DB 788 acttcgcgatatgttcatcttcagctacgacaaatgagtaataatcaaggtcatgtatgg 847
      |||||||
OY 496 cgggtagctcagttacatgattgtttaacccgttcagttcccaagactatccacc 555
      |||||||
```

Db 848 cgggtagtcagtcgattacagtggtttaacggtlcaagtlccccaagactlccacc 907
QY 556 cgttcgttcatcaaaaacatagaagtcagactcaggttgaaggttgctgtagag 615
Db 908 cgttcgttcatcaaaaacatagaagtcagactcaggttgaaggttgctgtagag 967
QY 616 ataacagctctccttgcttgatctcgatataccacaagaagtgtgtcaagaatgagt 675
Db 968 ataacagctctccttgcttgatctcgatataccacaagaagtgtgtcaagaatgagt 1027
QY 676 acgactggttggtgatactgtgatacgaactca----- 711
Db 1028 acgactggttggtgatactgtgatacgaactcaactcgaataatcttcccttca 1087
QY 712 -----ttgacggcttcgfatcgacaagta 737
Db 1088 caacttgctatcgatatacttaagaaatcagtltagcggctccctatcgacaagta 1147
QY 738 aaacagctccagaagactcttgcccggttacaacaagaagcgagcggtgactgtatc 797
Db 1148 aaacagctccagaagactcttgcccggttacaacaagaagcgagcggtgactgtatc 1207
QY 798 ggcgaggtgcttcgaagtgatcccggtcctaactgttccctacagaagtcataagcg 857
Db 1208 ggcgaggtgcttcgaagtgatcccggtcctaactgttccctacagaagtcataagcg 1267
QY 858 gtactgaactatcccat----- 874
Db 1268 gtactgaactatcccatgtatggttcctccaacatgagccttcttgcaagtcctatctc 1327
QY 875 -----ttactatccactcctcaagcgttcaagtlcaactccg 912
Db 1328 ctacgaagacggtaaacaggttactatccactcctcaagcgttcaagtlcaactccg 1387
QY 913 gcaagataggaagactcttacaacatgatacaaacgcgttaaatccgactgtccagactca 972
Db 1388 gcaagataggaagactcttacaacatgatacaaacgcgttaaatccgactgtccagactca 1447
QY 973 cactccttggaacatctgtcgagaacccagcaacccaggttgcctt----- 1021
Db 1448 cactccttggaacatctgtcgagaacccagcaacccaggttgccttgccttgccttcc 1507
QY 1022 -----ttaaac 1027
Db 1508 ctattattccgttcccaatttccacaacagaaacccactaacagaagcaagttaacac 1567
QY 1028 caagacataagccctcgccaagaagtcgcaagatcatcatcatcctcaagaagatccc 1087
Db 1568 caagacataagccctcgccaagaagtcgcaagatcatcatcatcctcaagaagatccc 1627
QY 1088 catcatctacgcgcgcgaagaacagcactacgcgcgcgcgaagaaccccgcaacgcgga 1147
Db 1628 catcatctacgcgcgcgaagaacagcactacgcgcgcgcgaagaaccccgcaacgcgga 1687
QY 1148 agcaaccttgcttcggtgtacccgacgcgacgcgagctgtatacaagttaattgtcctcgc 1207
Db 1688 agcaaccttgcttcggtgtacccgacgcgacgcgagctgtatacaagttaattgtcctcgc 1747
QY 1208 gaacgaatccggaactatgcatatagaacaagatacagaatcgttgactata----- 1260
Db 1748 gaacgaatccggaactatgcatatagaacaagatacagaatcgttgactataagtaag 1807
QY 1261 ----- 1260
Db 1808 cacaaccttaaccataatgacctatcttcagagtatctgaacaagaagactaat 1867
QY 1261 -----agaactggcccatctacaagaacgacacaagatcgccatgycgaag 1308
Db 1868 caactgcaatacagaactgcccacttacaagaacgacacaagatcgccatgycgaag 1927
QY 1309 gcaacgaatggtgctgcagatcgtgactatctgttcaacaagaagtgcttggtgagtctgt 1368
Db 1928 gcaacgaatggtgctgcagatcgtgactatctgttcaacaagaagtgcttggtgagtctgt 1987

QY 1369 ataccctctccttgagtggtgcgggttacacagccgcgaacgaattgacgagatcgtg 1428
Db 1988 ataccctctccttgagtggtgcgggttacacagccgcgaacgaattgacgagatcgtg 2047
QY 1429 gctgcagacccgtgacggttggttcgagatggaatgtgcctgttccctatgcaagtgggc 1488
Db 2048 gctgcagacccgtgacggttggttcgagatggaatgtgcctgttccctatgcaagtgggc 2107
QY 1489 taacctaggatgtratacgcgactggaaggttgccaggttagcaagatctgtaagctcgt 1548
Db 2108 taacctaggatgtratacgcgactggaaggttgccaggttagcaagatctgtaagctcgt 2167
QY 1549 gaaggtggaagatataatgatgtactgtactatcaactctgcatcgtgaacgtgattga 1608
Db 2168 gaaggtggaagatataatgatgtactgtactatcaactctgcatcgtgaacgtgattga 2227
QY 1609 gtttgatgatacagtttgagtgctgtactgtctcatcccttatactctgattgtttt 1668
Db 2228 gtttgatgatacagtttgagtgctgtactgtctcatcccttatactctgattgtttt 2287
QY 1669 cgaacctaatgccaagcagcgtagtctattatagaa 1707
Db 2288 cgaacctaatgccaagcagcgtagtctattatagaa 2326

RESULT 4
AAFI1252
ID AAFI1252 standard; cDNA, 1389 BP.
XX
AC AAFI1252;
XX
DT 13-MAR-2001 (first entry)
XX
DE Aspergillus niger EST SEQ ID NO:3775.
XX
KW Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus niger.
XX
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US07781.
XX
PR 22-MAR-1999; 99US-0273623.
PA (NOVO) NOVO NORDISK INC.
PA (NOVO) NOVO NORDISK AS.
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
DR WPI: 2000-594572/56.
XX
PT Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PR substrate of expressed sequence tags -
XX
PS Claim 87; Page 1694-1695; 3161pp; English.
XX
CC The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring

Sequence 1389 BP; 354 A; 384 C; 325 G; 326 T; 0 other;

Query Match 25.3%; Score 439; DB 21; Length 1389;
Best Local Similarity 98.9%; Pred. No. 2.2e-92;
Matches 442; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1174 ccacagcagagctgtacaagtaattgtcctccggaagcaatccggaactatgacatta 1233
DB 1364 CCGGCGTAGTGGCTGTCAAGTTATGTCCTCCGCCGAACCAATCGGAATATGCAATTA 1305
QY 1234 gcaaaatacaagatctgtgacctacaagaactggccactacaagaacgacacaaga 1293
DB 1304 GCAAAATATACAGATTCTGACCTACAAAGACTGGCCCATCTACAAGACGACACAAGA 1245
QY 1294 tcgcaatggcgaagcgaagaatgggtgcgaatcgtgtactcttgtccaagaaggtg 1353
DB 1244 TCGCCATGCGCAAGGCGACAGATGGGTCCAGATCGTACTATCTTGTCCAAAGGGTG 1185
QY 1354 ctccgggtgattcgtataccctctccttgagtgtgcgggttacacaagccgacgaat 1413
DB 1184 CTTCGGGTGATTCGTATACCTCTCTTCAAGTGTGCGGTTACACAGCCGCGCAAT 1125
QY 1414 tgcaggaagtcattggctgcacgacccgtgacggttggttcggaatggtgctgttc 1473
DB 1124 TGACGAGGCTCATTTGGCTGCACGACCGTACGCTTGCTCGGATGGAATGTGCTGTT 1065
QY 1474 ctatggcaggtggctgactaaggtatgtatccgactgaagaattggaggttaaga 1533
DB 1064 CTTATGCAAGTGGCTACTAGGATTTATTCGACTAGAAAGTTGGAGGTACCAAGA 1005
QY 1534 tctgtatagctcgttgaaaggtgagagataatgagtactgctatcaatctgcat 1593
DB 1004 TCTGTAGTAGCTCGTGAAGGTTGAGAGATATGATGTACTGATTCATCTGCAAT 945
QY 1594 ggaacgtgagtttgagtttgatgataca 1620
DB 944 GGACAGTGAATTGATTTGATGTACA 918

RESULT 7

AAQ25079 standard; DNA; 1404 BP.

AAQ25079;

17-NOV-1992 (first entry)

Alpha-amylase variant encoding leucine at position 84.

Mutant; maltose; malto-oligosaccharides; Saccharomycopsis; fibuligera;
polymerisation; DP; transglycosifier; ss.

Saccharomycopsis fibuligera.

Key Location/Qualifiers
329
mutation /*tag= a

/*note= "mutated to thymine"

JP04108386-A.

09-APR-1992.

28-AUG-1990; 90JP-0226112.

28-AUG-1990; 90JP-0226112.

(AGEN) AGENCY OF IND SCI &.

WPI; 1992-171652/21.

P-PSDB; AAR24136.

Variant alpha-amylase gene for mfr. of malto-oligosaccharide(s) -
is obtd. by mutating the nucleotide at position 329 of the
Saccharomycopsis fibuligera wild-type sequence to thymine.
Claim 1; Fig 1; 10pp; Japanese.

The variant alpha amylase gene was obtd. by mutating the 329th
nucleotide of the alpha amylase gene of Saccharomycopsis fibuligera
to T. This mutation results in substitution of the wild-type amino
acid at position 84 of alpha-amylase by leucine. The variant alpha-
amylase is high in transglycosylating activity. The variant alpha-
amylase may be used to prepare malto-oligosaccharides with a degree
of polymerisation (DP) of at least 7, by inversion of the malto-
oligosaccharide.

Sequence 1404 BP; 407 A; 271 C; 294 G; 432 T; 0 other;

Query Match 21.8%; Score 378.6; DB 13; Length 1404;
Best Local Similarity 55.7%; Pred. No. 2.3e-78;
Matches 754; Conservative 0; Mismatches 584; Indels 15; Gaps 1;

QY 121 ctgcgagctggcgaatccgaatccattatctccttcacaggaatgtaagagagc 180
DB 11 ctgaataatggagatacaagcttattatcaaatgtcacgtgacagattgttagaacg 70
QY 181 atggtcgcagactgcgacttgaatactcgcgatacgaataactgtgtggaacatgc 240
DB 71 atggtgatacaagtgcttcctcgttaacacagaagatgaacttactgtgtgtcttcc 130
QY 241 agggcattatcgacaaattgactatataccaggaagaggttcacagccatctggata 300
DB 131 aaggcattataaagaattgattatcaatcaagaatagtgcttactgtcatttgatt 190
QY 301 ccccggttacagccagctgcgccagacacacgacataatgagatgcttaccatgctact 360
DB 191 ctccagttgtgaaacattcccgatatacacagacatggttactatcatcagttact 250
QY 361 ggcagcaggtatatactctctgaacgaataacagcagcactgaagtactgaagaaggc 420
DB 251 ggaatgaagaacatauaaataatgaagaacttgactcgtcgtatgatttgaagctt 310
QY 421 tctctcgcccttcagaaaggggaggtatctatctatgctgcagtgtggttcaacata 480
DB 311 tggcaacaagaattgcacagatcgtgatatgttgaatggtgatcgttaccacacatt 370
QY 481 tggcctatgtagagcggtagctcagtgattacagtggtgtttaaaccgttactgcc 540
DB 371 acgacagtgatggcagtgagatagatcgaattaccacagagctacaccggttaacagcc 430
QY 541 aagactactccaccgcttctgttcttcaataactatgagaatgagctcaggtgag 600
DB 431 aaaaatcctccataactactgtccttattcaacaactatgagccaaagctcaggttcaaa 490
QY 601 attgctgtagagagataaacacgtctccttgctgcatcgtatccacacaaagatgtg 660
DB 491 gtgtcgtggaagtgactcttcagttgcatccagattgagaaggaagatagagag 550
QY 661 tcaagaatgaatgtagactgtgtggagatcattggtatccactactccatgaagcc 720
DB 551 tggcctcagtttcaattcttctgtttaaagaattgttggcaattactaatgtatgt 610
QY 721 tccgtatcgacacagtaaacacagctcgaagagactcttgcccggttaacaagaagcg 780
DB 611 taagaattgatagtcctaacaatgtagcaagacttctccggatattgttagtgcat 670
QY 781 caggcgttactgtalccgagagtgctcgaacggtatcccgctcactgtccctac 840
DB 671 ctgagatttactcagtagaggaagtttccaagagacacaggtataatgocctacc 730
QY 841 agaacgtcagagcgcgtactgaactatccattactatccactcctcaagcctta 900
DB 731 aaaaatactccaggggttagtaattatcattgactaaccacacagagattttta 790

QY	901	agctaaccttcg	gcagcatg	gaagcact	ctctaa	caactg	taca	caacg	gtca	aatccg	act	960
Db	791	aaactactgatt	caagttcc	agtgag	ttgta	ctaa	atgtat	ccaag	cgltgt	cttcag	tt	850
QY	961	gtccagactca	caactctt	ggtg	ccactct	gtgc	gtgag	agac	cgac	caac	ccag	1020
Db	851	gttcggtacc	aaactttgt	gtgaca	actttgt	gtaga	aatcc	gata	tcgata	tga	aaag	910
QY	1021	cttaccacca	gcgacat	agccct	ctgc	cca	agaag	tgc	gacat	tcata	ctcc	1080
Db	911	caatgacc	acgcg	ccaaga	gtttgat	ttcc	aatg	ctata	ttc	at	ttc	970
QY	1081	gaatccca	ctac	tacgc	ccg	gcga	agac	ga	ctac	tagc	cg	1140
Db	971	gtatctct	gtcatt	ctacta	ctat	gac	aga	gaac	aaag	cttg	gc	1030
QY	1141	accgcga	agac	acttgc	ctgc	gtgc	gtacc	cgac	cgac	gcag	ctgt	1200
Db	1031	acagagag	gccttgt	gtgtac	ctcc	gtac	aca	aga	gag	tgact	attaca	1090
QY	1201	ctctccg	gaac	gcga	atcc	ccga	actat	g	ccatt	agaca	agat	1260
Db	1091	ccaaag	ctaat	gtgc	gcga	aga	acgc	gcgt	ttat	ataa	gact	1150
QY	1261	agaactg	cccatc	ta	ca	aa	ga	gc	ac	ga	ca	1320
Db	1151	agcttct	gtgac	cttcca	atg	acac	ag	ctat	gtg	caac	aaag	1210
QY	1321	cgacag	atcgt	ga	ctat	ctgt	tcaca	aa	gg	tgtc	gtgc	1380
Db	1211	ctg-----	-----	tttca	aca	ac	ctt	gtgt	cc	ag	cg	1255
QY	1381	tgaagt	gtgc	gggtt	aca	gc	ccgc	gc	ga	ctg	ac	1440
Db	1256	tttcca	aca	ag	gttaca	g	atc	gc	gtg	ga	gatt	1315
QY	1441	tga	cgtt	gtgt	ccg	at	g	at	g	at	g	1473
Db	1316	ttag	cgc	gag	ctg	ctg	actt	aca	ag	ttt	ctat	1348

RESULT 8

AAAN70916

AAAN70916 standard; DNA; 4214 BP.

AAAN70916:

03-MAY-1991 (first entry)

Sequence encoding alpha-amyase from plasmid psf alpha 1.

Amyase; ds.

Saccharomyces fibuligera HUT7212.

Key Location/Qualifiers

CDS 1531..3015

/*tag= a

JF62104576-A.

15-MAY-1987.

31-OCT-1985; 85JP-0244892.

31-OCT-1985; 85JP-0244892.

(FUKU/) FUKUI S.

WPI; 1987-173694/75.

P-PSDB; AAP70571.

PT Amylaseprod. - comprises culturing microorganism transformed
 PT with vector deoxyribonucleic acid, accumulating and collecting
 XX amylase
 XX
 PS Disclosure: Fig 1; 14pp: Japanese.
 XX
 CC The plasmid may be used to transform an E.coli expression system for
 CC the stable production of amylase, useful in ethanol fermentation.
 CC See also AAN70917.
 XX
 SQ Sequence 4214 BP; 1249 A; 784 C; 860 G; 1321 T; 0 other;

Query Match	21.8%;	Score 377.4;	DB 8;	Length 4214;
Best Local Similarity	54.8%;	Pred. No. 6e-78;		
Matches 778; Conservative	0;	Mismatches 626;	Indels 15;	Gaps 1;

OY	55	tgctgcgctgtaggctctattctctctgacggccttcagctgcgcgcactgctttggctg	114
Db	1553	tgcttgcctcatctgctcgtcctctgtttatgctccaaccagtgactctattccaagaagaa	1612
OY	115	caacgcctgcgagctcgcgcgacgcacatccatttctctctccacggatcagatttgaa	174
Db	1613	ctaatgcgtataaatgagagatcacagctctatttccaattgtacagagatttgca	1672
OY	175	ggaacgattggtgcgcgcgactgcgcattgtaactgcgcgataagaatactggttgaa	234
Db	1673	gaaccgattggtatatacaagtgtctctctgtaacacagaagatagacttactgtgtglt	1732
OY	235	catgacaggaacatactgcagacagctggagctatactcagaggaatggcttcaacgacat	294
Db	1733	ctttccaagagctcctataagaagttggattacatcaaatatgvgcttactgttatlt	1792
OY	295	ggaatcaccccgcttacaagcccacgctgtcccacagaccacgcgatatagatgctacag	354
Db	1793	ggattcttcacgttgttgaacaacttcccgatacaacgatatgattgttactaatg	1852
OY	355	gctacttgcagcagatatatactctctcgaaacaaactaagcgcctcgaattgctga	414
Db	1853	gttactgtatgtaagaacatatacaaatatgaanaacttgcactgcgtgtattgta	1912
OY	415	aagcgcctctctcgcgcctctatgagaggatgatacttctgtgtcgatggtgtgca	474
Db	1913	agctcttgcacaaagaatattgcacgactcgtgatagattgttgaatgvtgatatacgttaca	1972
OY	475	accataatggtcatatgtaggcggtgtagctcagtcgattacagatgtgtttaaaccgttca	534
Db	1973	accattacgcagatgtagcagttgagatagatcatgattactcagagatnaaccocgltta	2032
OY	535	gttcccaagactacttccaaccgctctgttttcaatccaacatagaagatcagactcagg	594
Db	2033	acgaccacaaagtaactcataactactctgtctatttccaactatgataaccaagctcagg	2092
OY	595	ttgagattgctcgtctagagagataaacactgtctctctgctatctcgataccaccaagg	654
Db	2093	ttcaaatgttgcctgggaagagtgactccttcagttgcattaccagatttggaaacggagaata	2152
OY	655	atgtgttcataagaatgaaatgtagcactggtgtgagatattgtgtalccgaactccatgt	714
Db	2153	ggagacgttgcctcagtttccaattctctgtgttaagaatttgtgtgcaactaatgt	2212
OY	715	acgcgctccgatactgacacagtaaaaaacgttccagaagactctgcgcccggtataca	774
Db	2213	atgtgttitaagaattgatagtgtctaacaactgtgacaaagcgttltcccgagtttgtta	2272
OY	775	aagccgcagcgtgtgactgtatcgcgcgaggtgctcgcagcgtgtatccgcgcctaactgtc	834
Db	2273	gtgcatctcgagtttactcagcagtgagcgaagtttccaagaggaccagctataatgtcc	2332
OY	835	cctaccagaacgtcatagacggcgtactgtgaactatcccaattactatatacactctcaag	894
Db	2333	cataccaaatatcactccagaggtgtagtaattatcttcaacttgcattaccacaaacgaat	2392

QY	895	ccttcaagttcaacctccgcgagatgagaaagactcttacaacttgatacaacgcgttcaat	954
Db	2393	ttttaaactactgattcaatgaattccaaagttagttgactcaaatgttatccaagcgtttct	2452
QY	955	ccgactgttccagatccaacacctccttgcgcacatttgcagacaacgaaccaacgaagt	1014
Db	2453	ccgaatttccgatccaactttttgacaaactttgtagaanaatacagataatgaagaagt	2512
QY	1015	tgcgttttccacccaacagacatagccctcgcgaagaagctgcagacatcatcactca	1074
Db	2513	tgcgttccatgacccaagacgaacaaatttgattcttaatctatgcatcttcctcttgg	2572
QY	1075	acgcgcggaatcccatatcatcttaacgcgcgcgaagaagaagcactacgcgcgcggaagcgc	1134
Db	2573	gttagttgattccgttacttactatctagaaagaagaagctttgcgcgaaanaagttgc	2632
QY	1135	ccgcgaacccgcgaagcaacactgtgctctgcgcaccgcgcgcacgcgcgcgcgtctgaagt	1194
Db	2633	caaacacaagagagagccttctgtgtatccgcgtctacacaagaagagttgactttcaagc	2692
QY	1195	taattgctctcgcgcgaacgcacatccgcgaactatgtccattgacaagaatacagatctcgtga	1254
Db	2693	tcattgcacaagctaatgctgcgcgaagaagccgcgctttatcaagatcaagctatgcga	2752
QY	1255	ccctacaagaactggtcccatctacaagaagcacaacacgctgcgcctgcgcgaagggcacag	1314
Db	2753	ccctgcgcagcttctgtgattctttccaatgacatgttatgtgacaaaagaagggcacg	2812
QY	1315	atgggtcgcagatcgtgactactctgtccacaagaagtgactccgcgtgattcgttatccc	1374
Db	2813	ttgtttctgt-----ttttcaacaacttggtccagcggttcttctgtg	2857
QY	1375	tctccttgatgtagtgcggtgtacacacagccgcgcgaagcaattgacgcgaggttcattgctgcga	1434
Db	2858	tgactatttccaacaacaggttaccagttccggtgagagatttgtagaagtttgacatgcga	2917
QY	1435	cgaccggtgacggtgtgttcggtacggaagagtagtcctgttc	1473
Db	2918	gtactgttagcgcgacgtctgcacttacaattttctatcc	2956

RESULT	9
AA077668	
ID	AA077668 standard; DNA; 1404 BP.
XX	
AC	AA077668;
XX	
DT	16-JUN-1995 (first entry)
XX	
DE	Variant alpha amylase deriv. from Saccharomycopsis fibuligera.
XX	
KW	alpha amylase; carbohydrate hydrolase; increased activity;
KW	lysosine residue; enzyme centre; mass production; oligosaccharide;
KW	variant; cyclomaltohextrin glucanotransferase; ds.
XX	
OS	Saccharomycopsis fibuligera.
XX	
FT	Key
FT	misc_difference 247..249
FT	location/Qualifiers
FT	/*tag= a
FT	/note= "the wild type sequence TAY was mutated to AAC to give a variant enzyme"
XX	
PN	JP06253836-A.
XX	
PD	13-SEP-1994.
XX	
EP	04-MAR-1993; 93TP-0065303.
XX	
PR	04-MAR-1993; 93TP-0065303.
XX	
CA	(AGEN) AGENCY OF IND SCI & TECHNOLOGY.
XX	

DR WP1: 1994-338987/41.
DR P-PSDB; AAR63187.
XX
PT Variant carbohydrate hydrolase(s) with increased activity -
PT consists of e.g. alpha-amylase with tyrosine residue in enzyme
PT centre, useful for mass-prodn. of oligosaccharide(s)
XX
XX
PS
XX
XX
CC AAO77665-8 encode variant alpha amylases, composed by substituting
CC bases 247-249 of the structural gene region, with TTC, TGG, CTC or
CC AAC. These substitutions result in the 83rd amino acid residue
CC (tyrosine) of the wild type sequence being changed to phenylalanine,
CC tryptophan, leucine or asparagine respectively. The substituted
CC amino acid is present in the active site of the enzyme and confers
CC increased activity on the enzyme. The variants are useful for the
CC mass production of oligosaccharides. (see AAO77666 for the variant
CC structure of a cyclomalto-dextrin glucoamyltransferase).
XX
SQ Sequence 1404 BP; 408 A; 272 C; 293 G; 431 T; 0 other;

Query Match	21.6%	Score 375.4	DB: 15	Length 1404	
Best Local Similarity	55.6%	Pred. No. 1.3e-77			
Matches 752	Conservative	0	Mismatches 586	Indels 15	Gaps 1
QY 121	ctgcgcagctcgcgcacgcgcacccatccattcttctctccctccacgcatcgatttcgaagacgcg	180			
DB 11	ctgataaatgtgagatccacgctctattctatccaattgtcactgcagacgattgtctgaaccg	70			
QY 181	atggctgcgcagctcgcgacttgcgtttaactcgtcgcatcagaanaabactcgtgtgtggaacatgcg	240			
DB 71	atgtgtatacaacgctctcctctgaacacagaagatagacttactcgtgtgtctcttc	130			
QY 241	aggcgcatccgcacgaagtgtgactatataccgaagatgtggtctccacgcacatcgatca	300			
DB 131	aaggcatcatanaaagaatttggattatatacaagaagatagtggtcttaccgtctatttggatt	190			
QY 301	ccccgtttacagcccgacgtcgcccgacacgcgcatactggaatgcctaccatgtgcact	360			
DB 191	ctccagttgttggaaaaaatctcccgataaacacgcgcatatgtgatacttatcatgtgtact	250			
QY 361	ggcagcagagatatactctctgcacgnaaacatcagcgcactgcagatgaatttgaagcgc	420			
DB 251	ggaatgaagaacatatacaataaattatgaaacacttgcgtactgcgtatgttgaagctct	310			
QY 421	tctcttcgcgccttcatacgagaggggagatcatctatgctcgatgtgtgttgcatacata	480			
DB 311	tggacacaagaatttcgcgcgactcgtgatatgttctaagtggtatatactgtaccacaacatt	370			
QY 481	tggcgctatgatggagcggtgagctcagctcagctatcacagctgtgtttaaacgcgttcagctcc	540			
DB 371	acgcgcattgtagtgcaggtggagagatagatcatcgattactccagagtgacaccccgcttaacgacc	430			
QY 541	aaagactatccaccgcgtctctgttccatccaactatagaatgaatcagaactcaggttgaag	600			
DB 431	aaaagtacttccataactactcgtctatcttccaactatgatgaccacaagctcaagttcaaa	490			
QY 601	attcgtcgtcagtaggagataaacactgctctccttcgcgcgactcgcgatatccaccagaagtgtg	660			
DB 491	gttcctcggaaagtgactcttcactcagttgcattccagatttgaagacggaagatgcgacg	550			
QY 661	tcaagaatgaatgtgtacactgtgtgggagatcatgtgtatcgcgaactccacatcgaagcc	720			
DB 551	tggcctcagtttcaattcttgggtttaaagatttgggttgcacatctaccatctgaatgtgt	610			
QY 721	tccglatcgcacacglaanaaacgcctccagaagagacttcgcgcgcggtgtacaaacaaagcg	780			
DB 611	taagaattgtagtgcgttaaacacatgttgagaccagcttttcccggaatttgttagtgcatt	670			
QY 781	cagcgcgttactgtatctcgcgcgaggtgctcgcagcgtgtatccgcgcctacactgttccctaac	840			
DB 671	ctgagattactcagttcggaggaatttcccaaggagacccacattatcatcagcccatcatt	730			


```

QY 841 agaagctatgagcggtactgaactatccattactatccactcctcaagccctca 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 731 aaatattatccaggggttagtaattcatctatccatcccaacagcagatttlla 790
QY 901 agtcaactccgagcagctgacagcctctacaacatgacacacgctcaatccgact 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 791 aaactactgattcaagtccagtgagtgactcaaatgatttcaagcgttgcctccggt 850
QY 961 gtccagactcaacactcctgggacacatcgtcgaagaccagacacccacggttcgt 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 851 gtccgactcaacttgttgcaaacatttgaagaatacagataatgaaggttcgtt 910
QY 1021 cttaaccaagacatagcctcctcgcaagacgctcgacgttatactatcctcaagacg 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 911 caatgacagcagcacaagatttgatttctatgtattgtatcttcttgggtgatg 970
QY 1081 gaatcccatcatctacgcccgaagacagacactacgcccgcggaagcagcccgga 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 971 gtattccgtgactattactatgacagacaaagccttgagcggaaaagtgacccaaca 1030
QY 1141 accgcgaagcaacctgctcgggctacccgacgcagcagctgtacaaagttaattg 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1031 acagagagccttgygtatccgctacacaagaagagtgactattacagactcatg 1090
QY 1201 cctcgcggaagcaatccggaactatgcatctagcaagaagatcgttgacctaca 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1091 ccaaggtcaatgctgcggaagacgcgcgttctatcaagactcaagctaatgcccctgcg 1150
QY 1261 agaactgcccatactacaagaacacacaaagatcgccatcgcaagggcacaagtgtg 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1151 agcttcctgtatctcttccaatgaccatgttatgacaaaagaagcgagcgttctt 1210
QY 1321 cgcagatcggtactactctgtccacaagaaggtgctcggtgtgattcgtatccctcc 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1211 ctg-----tttccaacaacctgtgttccagcgtctctctgaatgacta 1255
QY 1381 tgagtggtgcggttacacagccgcgcgaagtgaagctcaattgctgcagcagcgcg 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1256 ttccaacaacaggttacaggttccggtgaggttggtagaagtttgcatagcagttactg 1315
QY 1441 tgacggttggttcggaatggaatgtgcctgttc 1473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1316 tttagcgagcagctctgactacaagttctatccc 1348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
AAQ77665
ID AAQ77665 standard; DNA; 1404 BP.
XX
AC AAQ77665;
XX
DT 16-JUN-1995 (first entry)
XX
DE Variant alpha amylase deriv. from Saccharomyopsis fibuligera.
XX
KW alpha amylase; carbohydrate hydrolase; increased activity;
XX tyrosine residue; enzyme centre; mass production; oligosaccharide;
XX variant; cyclomaltoextrin glucanotransferase; ds.
XX
OS Saccharomycopsis fibuligera.
XX
FH Key Location/Qualifiers
FT misc_difference 247..249
FT /tag= a
FT /note= "The wild type sequence TAY was mutated to
TTT to give a variant enzyme"
XX
PN JP06253836-A.
XX
PD 13-SEP-1994.
XX
PF 04-MAR-1993; 93JP-0069303.

```

```

XX
PR 04-MAR-1993; 93JP-0069303.
XX
PA (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
DR WPI: 1994-328987/41.
XX P-PSDB; AAR63184.
XX
PT Variant carbohydrate hydrolase(s) with increased activity -
PT consists of e.g. alpha-amylase with tyrosine residue in enzyme
PT centre, useful for mass-prodn. of oligosaccharide(s)
XX
PS Example 1; Page 13-15; 27pp; Japanese.
XX
CC AAQ77665-8 encode variant alpha amylases, composed by substituting
CC bases 247-249 of the structural gene region, with TTC, TCG, CTC or
CC AAC. These substitutions result in the 83rd amino acid residue
CC (tyrosine) of the wild type sequence being changed to phenylalanine,
CC tryptophan, leucine or asparagine respectively. The substituted
CC amino acid is present in the active site of the enzyme and confers
CC increased activity on the enzyme. The variants are useful for the
CC mass production of oligosaccharides. (see AAQ77669 for the variant
CC structure of a cyclomaltoextrin glucanotransferase).
XX
SQ Sequence 1404 BP; 406 A; 272 C; 293 G; 433 T; 0 other;

Query Match 21.6%; Score 375.4; DB 15; Length 1404;
Best Local Similarity 55.6%; Pred. No. 1.3e-77;
Matches 752; Conservative 0; Mismatches 586; Indels 15; Gaps 1;

QY 121 ctgcgagctgcgacgcgaatccatctattctctctcacaagcagctattgcaagcag 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11 ctgataatggaatgatacagctctattatcaaatgtgtaactgacaaatttgacagcgg 70
QY 181 atgggtgcagcagctgcgactgttaatactgcggaatgaataactgttggtgaacatgyc 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 71 atggtatacaagtgctctcctgaacaagaagatagacttactgtgtgtctcttccc 130
QY 241 agggcatcatcgcaagttgactatataccaggaatgggtcttaccatctgtgata 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 131 aaggcatcacaagaagaattgattacatcaagaatggcttactgtcatttgaatt 190
QY 301 ccccgcttaagcccgctgcgccagacacccagacatagagatgctactcaatgact 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 191 ctccagttgttgaanaacattcccgataacacagcatatggttatcgttatacagttct 250
QY 361 ggcagcagatatatactctctgaacgaaactacgagcactgcagatgactgaaagcgc 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 251 ggaatgaagacatatatacaaaattatgaaaacttgtgactgctgatttgaagctt 310
QY 421 tctctcgccctcatgagagagggtgattatcttatagtcgtatggtgtgataccata 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 311 tggacaagaattgcaagacgctgatatgttgaatggtgatatcogttaccaacatt 370
QY 481 tgggtatgataggcgggtagctcagctcagctacagtggtttaaacogtcaattcc 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 371 acggcagtgatgagagtgagatagatcgtattacagaggtacacccgcttcaagacc 430
QY 541 aagactactcaaccggttctgttcaatcaaaaactatgaaagatcagactcaagttgag 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 431 aaaagtacttccataactactgttatttcaactatgtagcacaagctcaagttcca 490
QY 601 atgtgctagtgagataaactgtctcccttgctgactcgcgtatccacccaagaatggtg 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 491 gtgtctggaagtgactcttcagttgcatltaacagattggaagaagaaatagagag 550
QY 661 tcaagaatgaaatgtacagctgggtggatcaatgtgatalcgaactactccattgacggcc 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 551 tggcctcagtttcaattcttgggttaagaatttgttgccaattactcactgaattggtt 610
QY 721 tccgtatcgaacaggttaaacacagctcagaagaagactctgcccgggttaacaaagccg 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

QY 1601 gagttcagtttgatgtagcagcttggaagcttactgctgacacccctataacttcga 1660
    |||||||
DB 360 aagtttgagtttgatgtagcagcttggaagcttactgctgacacccctataacttcga 418
    |||||||
QY 1661 tggtttcgagcccaatgccaaagcgcgtagctgctatcttaagaagaaaaaa 1720
    ||| | ||||||| ||||||| ||| | |||
DB 419 ----gtttcgaaactaatgcaa-cacgctagtctctatagaagaagmctcgattaa 473
    ||| | ||||||| ||||||| ||| | |||
QY 1721 aaaaaa 1727
    |||||||
DB 474 aaaaaa 480

RESULT 12
AAQ77667
ID AAQ77667 standard; DNA; 1404 BP.
AC
AAQ77667;
DT
16-JUN-1995 (first entry)
DE
Variant alpha amylase deriv. from Saccharomycopsis fibuligera.
XX
KW alpha amylase; carbohydrate hydrolase; increased activity;
KW tyrosine residue; enzyme centre; mass production; oligosaccharide;
KW variant; cyclomaltohextrin glucanotransferase; ds.
XX
OS Saccharomycopsis fibuligera.
XX
FH Key Location/Qualifiers
FT misc_difference 247..249
FT /*tag= A
FT /note= "the wild type sequence TAY was mutated to
FT CTC to give a variant enzyme"
XX
PM JP06253636-A.
XX
PD 13-SEP-1994.
XX
PF 04-MAR-1993; 93JP-0069303.
XX
PR 04-MAR-1993; 93JP-0069303.
XX
PA (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
DR WPI; 1994-328987/41.
XX
P-PSDB; AAR63186.
XX
PT Variant carbohydrate hydrolase(s) with increased activity -
PT consists of e.g. alpha-amylase with tyrosine residue in enzyme
PT centre, useful for mass-prodn. of oligosaccharide(s)
XX
PS
XX
Example 1; Page 18-20; 27pp; Japanese.
XX
AAQ77665-8 encode variant alpha amylases, composed by substituting
CC bases 247-249 of the structural gene region, with TGC, TGG, CTC or
CC AAC. These substitutions result in the 83rd amino acid residue
CC (tyrosine) of the wild type sequence being changed to phenylalanine,
CC (tryptophan, leucine or asparagine respectively. The substituted
CC amino acid is present in the active site of the enzyme and confers
CC increased activity on the enzyme. The variants are useful for the
CC mass production of oligosaccharides. (see AAQ77669 for the variant
CC structure of a cyclomaltohextrin glucanotransferase).
XX
SQ Sequence 1404 BP; 406 A; 273 C; 293 G; 432 T; 0 other;

Query Match 21.6%; Score 373.8; DB 15; Length 1404;
Best Local Similarity 55.5%; Pred. No. 3, 1e-77;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

121 ctgcgacgtggcgatcgcaatcattattctcttcacgagatgatttgcagagcgg 180
||| ||||| ||||| || ||||| || ||||| || ||||| ||||| ||||| |||||

```

Db	11	ctgttaaatggaatcacagctctatttcaaatgttcaactgtacagattgtgtagaacg	70
Qy	181	atggctgcagactgcgactgtgtaatactgcgatactgcagataactcgtgtggaacaatggc	240
Db	71	atgggtgatacaatgtcttctctgttaacacagaagaataagcttactgtggtgtcttcgc	130
Qy	241	aaggcatcatcgcgaaggttggactatataccagggaatvggcttcaacagccatctgtatac	300
Db	131	aagcgcataaagaagatttggattacatcaagaatavgtgcgttaccgtcatcttggattt	190
Qy	301	ccccgttaccgcccgcgtgcgccacgacccgcgtatggagatggcttaccatgctgcact	360
Db	191	ctccagattgttgaanaaatctcccgataacacagcatctgttatagtcttatacavgtctct	250
Qy	361	ggcagcgcgatatatactctctgaacgaanaactaagcagctgcagatgacttgaagcgcg	420
Db	251	ggatgaagaacaataataaanaattatgaaactcttggctacgtcgtatgatttgaagcgtt	310
Qy	421	tctcttcgcccctcatctatgagaagggtatcattatgctcgaatgtgttgggttgcatacata	480
Db	311	tggcaccaagaattgcacgatacgtataatgttgaatgttgatactgttaccacatc	370
Qy	481	tgggtatgatagtggcggtgtagctcagctcagattacaagtggttttaacagtttcc	540
Db	371	acgggaatgtagtggcagtgtagagatagtatcagatctaccagatgtacaccccggttcaagcc	430
Qy	541	aagactactctcacccgcttctgttctatccaataactatgaatcagatcagattgag	600
Db	431	aaagtaactctcataactactgtcttattccaactatgatgacacagctcagttcaaa	490
Qy	601	atttgtctgtagagataaacaactgtctctctgcgtatctcgtatataccccaagatgtg	660
Db	491	gttctgtggaaggtgactcttccagtttgcattacaagatttgaacgcgaagatagcgacg	550
Qy	661	tcaagaatgaatgtgtacgactggtgtagatcattgttatcgcgaactcccatctagcgcc	720
Db	551	tggcctcagtttcaattcttgggttttaagaattttgttgcataattaccatgtatggtt	610
Qy	721	tccgatactgcacacgtlaaacaacgtccacagaagactctgcgcgggtatacaacaagccg	780
Db	611	taagaattgtatagtgctaaacaatgtgacaagcgttttcccgattttgttagtgat	670
Qy	781	cagcgctgtactgtatcggcgaggtgtcgtcgaacgttatccggcctacactgttccctacc	840
Db	671	ctggagtttaactcagtagcggaagttttccaagagacccagctatatacatgtccatacc	730
Qy	841	agaaagtcagtagcgggtactgtaacactccattacatccatccactcccaacgctca	900
Db	731	aaatatacttccaaagggtatgtatattatccattgtactacccaacacgaagattttta	790
Qy	901	agtcacaactccgcgcgacatgtagcagccttacaacatgatacaacacgctcaaatccgact	960
Db	791	aaactactgtattcaagttcccgatgtatgtactcaaatgataatttcaagcgttcttccagtt	850
Qy	961	gttccagatctcaacactccttgggcacattcgttcgagaacacgaacccaacccagttcgttt	1020
Db	851	gttccgataccaacttgttgaacaacttvtganaaatacscgataatgaaagtccgtct	910
Qy	1021	cttaccacacacataagccctgcgaagaacgctgcagatcatcacaactccacaagcg	1080
Db	911	caatgacacagcgaacaaagtttatttcttaatgtatgtatgtattgttccctttgggtgagt	970
Qy	1081	gaatccccaatactacgacggcgacaagacagcatcgcgcggcggaaagacacccgcga	1140
Db	971	gtattctcttacttactatgtacagaagaacagctgttagcggaaaaaagtgtgacaaaca	1030
Qy	1141	accgcgaaagaaactggtctcttcgggtctaccgcgacccgacacgagcgtgtacaagttaatg	1200
Db	1031	acagaaagagccttgtgtatctcgcgtacaaacaagaaatgtgactattacaagcttacttg	1090
Qy	1201	cttccgcgaagcgaatcccggaatactgcatctagcagaataagaaatcgttgcatacta	1260
Db	1091	ccaagctaatgtctgcagaaacgcgcgcgtttatacgaagctcaagctatgtaccactgcg	1150

QY 1261 agaacgtgcccacatacaaacagacagatcgccacagggcacagatgagt 1320
 Db 1151 agcttctgtgactcttcaaatgacacatgacacaaagagcagcgttgctt 1210
 QY 1321 cgcagatcgtagactctgtccaaagaagtgctcgggtgattcgtataccctcct 1380
 Db 1211 ctg-----tttcaacaaccttgctccagcgctctctctgtgacta 1255
 QY 1381 tgaagtgtcggttacacagccgacagcaatgacagaggtcaattgctcagcagc 1440
 Db 1256 ttccacaacaagttacagctccggtgaggttgtagaagtttgatcagtcagtlacty 1315
 QY 1441 tgaaggtgtgtcggatgagaaatgtgcctgttc 1473
 Db 1316 tttagcgagcagctctgacttacaagttctcatcc 1348

RESULT 13

AA077666
 ID AA077666 standard; DNA; 1404 BP.

XX AA077666;

DT 16-JUN-1995 (first entry)

XX Variant alpha amylase deriv. from Saccharomyopsis fibuligera.

XX alpha amylase; carbohydrate hydrolase; increased activity;

KW tyrosine residue; enzyme centre; mass production; oligosaccharide;

XX variant; cyclomaltohextrin glucanotransferase; ds.

OS Saccharomyopsis fibuligera.

XX Key location/Qualifiers

FT misc_difference 247..249 /*tag a

FT /*note= "The wild type sequence TAY was mutated to

XX JP06253836-A.

XX 13-SEP-1994.

XX 04-MAR-1993; 93JP-0069303.

XX 04-MAR-1993; 93JP-0069303.

XX (AGEN) AGENCY OF IND SCI & TECHNOLOGY.

XX WPI: 1994-328987/41.

XX P-PSDB: AAR63185.

XX Variant carbohydrate hydrolase(s) with increased activity -

XX consists of e.g. alpha-amylase with tyrosine residue in enzyme

XX centre, useful for mass-prodn. of oligosaccharide(s)

XX Example 1: Page 15-17; 27pp; Japanese.

XX AA077665-8 encode variant alpha amylases, composed by substituting

XX bases 247-249 of the structural gene region, with TMC, TGG, CMC or

XX AAC. These substitutions result in the 83rd amino acid residue

XX (tyrosine) of the wild type sequence being changed to phenylalanine,

XX amino acid is present in asparagine respectively. The substituted

XX increased activity on the enzyme. The variants are useful for the

XX mass production of oligosaccharides. (See AA077669 for the variant

XX structure of a cyclomaltohextrin glucanotransferase).

Best Local Similarity 55.5%; Pred. No. 3..1e-77;

Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 ctgcgacgtgcgatacgcgaatccattatctccctccacggatcgattgcaagagc 180
 Db 11 ctgtaaatgtagatccacagctctattataaattgccaactgacaagattgctcagaacg 70
 QY 181 atggtgcgaactgtgacttgttaatacttcggatcagaataactgtgtgtaacatgac 240
 Db 71 atggtgataaagtgtctccgtgtaacacagaagatgacttactgtgtgtctctcc 130
 QY 241 agggcatcatgacaagtgtgactatataccaggaatagggcttccacagccatctgatac 300
 Db 131 aaggcacaataaagaatgtgattgatacaaaatagatagggcttactgtcatttggattt 190
 QY 301 ccccggttaacagccagctgcgcacagacccagacatctgagatgagtgcttaccagctact 360
 Db 191 ctccagttgttgaataaataatcccgataacacagacataatgttcttatacagtgttgc 250
 QY 361 ggcagcagagatatactctctgaacgaanaaactacggcactgacagtgtgaagcgac 420
 Db 251 ggaatgaagaacataatacaaaatlaatgaanaaacttgtactgtgattgtaagttct 310
 QY 421 tctcttcgccccttcattgagaggggagatattatgtgtcagtggtgtgtaaccata 480
 Db 311 tggcacaagaattgacagatcgcgatalgtgttaatggtgagatcgttaccacacatt 370
 QY 481 tggcgtatgataggagcgggtgagtcagtcagtcagtcagtcagtcagtcagtcagtc 540
 Db 371 acgcagagtgagcagtgagtgagatagatagatagatagatagatagatagatagatag 430
 QY 541 aagactacttccaccgcttctgttccatccaanaactatgaagaactagactagactgtgag 600
 Db 431 aagaacttccatccaactctctcttccatccaactatgaagaactagactagactgtgag 490
 QY 601 attgcgtgtagagataaacaactgtctcttcgttcgtatcgtatcgtatcgtatcgtatc 660
 Db 491 gtgcctgggaagtgagctctcttcagtgatcagatcagatcagatcagatcagatcag 550
 QY 661 tcaagaattgaattgtagcagctgggtggaatcattgtgatacgaactcattcagtcagcc 720
 Db 551 tggcctcagtttccaattctctgtgttaagaatttgttgcaattactcattcagttgtgt 610
 QY 721 tccgtatcgacaacagaacagcgtccagaagagactctgcggcggtacaaacagccg 780
 Db 611 taagaattgtagtgcctaaacatgtagcaagagcgttctccggatctgtgttagtgat 670
 QY 781 caagcgttactgtatcgcgcagagtgctgcagcgtgatacgcgtcactacactgtccctac 840
 Db 671 ctggaatttactcagtagagcgaagtttccaagagaccagcagctatataatgccatacc 730
 QY 841 agaagctgtagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900
 Db 731 aaaaattacatccaggggttagtaattacatcgttactacccaacacagagattttta 790
 QY 901 agtcaacctccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
 Db 791 aaactactgattcagatcagtgagtggtgactcaaatgatttcaagcgtgtgtccaggt 850
 QY 961 gtccagactcaaacactcctgggcacatcgttcgagaacacagcagcagcagcagcagc 1020
 Db 851 gtctggtatccaacttgtgacaacttgttagaataatcagataatgaaggttgcgtt 910
 QY 1021 cttaacacagaacagatagccctgcgaagaacgctgcagcagcttcatatcctcaagagc 1080
 Db 911 caatgacagcagcacaagatttgatttccatgcatgtattgtcccttgggtgattg 970
 QY 1081 gaatcccatcatcagcgcgcgcagcagcagcagcagcagcagcagcagcagcagcagc 1140
 Db 971 gttactcgttacttactatgacagagaagaagcgttgagcggaanaagtgacccaaca 1030
 QY 1141 accgcgaagcaacctggtctcgggtctacccgcagcagcagcagcagcagcagcagcag 1200

Query Match 21.6%; Score 373.8; DB 15; Length 1404;

Db	1031	acagagagcgcccttggtgtatccggtcacacaagaagagtgactattacaagctcaatg	1030
Qy	1201	cctccggaacgcacatccggaaactatgcatatgacaagaatacaagatctgtgaccta	1260
Db	1091	ccaagctaatgtgtccagaacgcgcgcgcttatccaagactcaagctatgcacactgcg	1150
Qy	1261	agaactgtgccatctacaagaagacacaacagatcgccatcgctgcaaggcagaagtggt	1320
Db	1151	agcttctgtgactcttccaatgacacacagttatgtacaacaagaagcgacgcttggt	1210
Qy	1321	cgcagatcgtgactactctgtccacaagaagtgcttcgggtgattcgtatataccctctct	1380
Db	1211	ctg-----ttttcaacaacctgtgtccgcgcgctctctctctgtatgacta	1255
Qy	1381	tgaatgtgtcggtttacaacagccgcgcagaatatgtgacgaagtcattgtctcacacgcg	1440
Db	1256	tttccaacaacgggttaaaagttccggtggaagatttggttaagaatttggaaatgcagaactg	1315
Qy	1441	tgaacggttggttcgcatggaatagtgcctgttc	1473
Db	1316	ttacgcgacgctctgacttacaagttctctacc	1348

RESULT	14	
ID	AAQ06388	
XX	AAQ06388 standard: DNA; 2289 BP.	
XX		
AC	AAQ06388;	
XX		
DT	02-FEB-1991 (first entry)	
XX		
DE	AMY1 gene encoding alpha-amylase.	
XX		
XX	Schwanniomycetes yeast cells; expression cassette; alpha-amylase	
KW	marker gene; regulon; signal peptide; terminator;	
KW	autonomously replicating sequence; ss.	
XX		
OS	Schwanniomycetes occidentalis.	
XX		
FH	Key	Location/Qualifiers
FH	CDS	550..2085
FT		/*tag= a
FT		/product=alpha-amylase
FT		508..520
FT	misc_signal	
FT		/*tag= b
FT		/function=transcription_initiation_site
FT	terminator	2086..2088
FT		/*tag= c
FT		/function=signal
FT	terminator	2180..2184
FT		/*tag= d
FT		/function=signal
FT	terminator	2231..2233
FT		/*tag= e
FT		/function=signal
XX		
PN	EP94538-A.	
XX		
PD	31-OCT-1990.	
XX		
PF	28-APR-1989;	89EP-0107780.
XX		
PR	28-APR-1989;	89EP-0107780.
XX		
PA	(RHEI) RHEIN BIOTECH GES.	
XX		
PI	Hollenberg C, Strasser A;	
XX		
DR	WP1: 1990-328670/44.	
DR	P-PsDB; AAR07574.	
PT	Transformed Schwanniomycetes yeast cells - contg. an expression	
PT	cassette contg. regulon, DNA coding for foreign protein and	

PT	terminator
PS	Disclosure, Fig 1B(1-6); 59pp: English.
XX	
CC	The expression cassette may contain a regulon (R), a signal peptide
CC	sequence (S), a foreign protein sequence (F) and a terminator (T)
CC	of which R and/or S and/or T are derived from the alpha-amylase gene
CC	of <i>S. occidentalis</i> .
CC	R is pref. a 1.8 kb BglII-XhoI fragment (bases -1 to -540 of this
CC	sequence).
CC	S contains all or part of one of the sequences represented
CC	in AA006383-87.
CC	F is e.g. a cellulase, interleukin, insulin-like-growth factor,
CC	interferon etc.
CC	T is pref. all or part of the terminator of this sequence
CC	(bp 1537-1740).
CC	AM1 acts as selective marker for the vector carrying the expression
CC	cassette.
CC	See also AA006389.
XX	
SQ	Sequence 2289 BP; 693 A; 398 C; 424 G; 774 T; 0 other;

Query Match	21.3%;	Score 369.2;	DB 11;	Length 2289;
Best Local Similarity	54.6%;	Pred. No. 4,1e-76;		
Matches 759;	Conservative	0;	Mismatches 628;	Indels 3; Gaps 1
Qy 121	ctgcgactcgcgatcgcgcatccgcatcttattctccctccgcgcatcgatttgcgaagcgg	180		
Db 665	cgatataaggaagaagccatcgattatccaatcgcttaactagatggttgcgaatctg	724		
Qy 181	atgggttcgacgactcgcgacttgaatactcgcggaatcagaatactctgtgtggaacttgc	240		
Db 725	atggcctcgagccaaagctgcgtgtttagtgaatgcgaactcgttggtagctctata	784		
Qy 241	aggcgactcgcgcaagcttggaactatattccgaaggaatgggttcacgaagcatctgatca	300		
Db 785	aagggaattcgcgaaggttggaattatattccaagtattggtttcacctgcgactgatct	844		
Qy 301	ccccgcttaacagcccgactgcgcccagacccagcatatggagatgcctaccatgctact	360		
Db 845	ccccagttgttgcgcaaatccctgcgaatactcgttatggtttagtcttaaccatggttatt	904		
Qy 361	ggcagcaggaatatactctctcgaagaaacgcgcaactgagatgacttgaagcgc	420		
Db 905	ggatgaaatactatgatgaattgaacactaatttggtaaccgttgatgaaatgaacaat	964		
Qy 421	tcctctcgcgccttcctcgaagagggggaatgatcttaattgttcgaatgtgtgtcctacata	480		
Db 965	tagctagcgatctgcattccacgaagcattgtattgatgtgcgagcttgtttaacaacatt	1024		
Qy 481	tgggcattgatgagcggggtagctcagtcgattacagtggtgttaaacggttaagctcc	540		
Db 1025	atgcttgaagcggagatggtgttcaagcgtatgattattctagttccaactccaatcaaac	1084		
Qy 541	aagactactccacccgctctggtttcattcaaaactatgaaatgcagactcaggtttag	600		
Db 1085	aatcttacttccacgattattgtttgttacaataatataatgatcaacaacaaatgttag	1144		
Qy 601	attgtgcctaggaagataacacatgtctccttgctgcgtatctcgataccacaaggaatgtg	660		
Db 1145	attgttggaaggtgatactcgaatgcctcccttccagattaaagtaaccgaggaataag	1204		
Qy 661	tcaagaatgaaatgttacgactgggttggaatcattgtgatccgaactatccattcaagcc	720		
Db 1205	ttataaggaattatccaacttgggtgttcagatttgttccaacactatccaatcgatgtct	1264		
Qy 721	tcgctatcgacacagtaaaaacgctccaagaagactctcggccgggtacaacaagcgc	780		
Db 1265	taagaattgatagtgcaaacgactagataccggttcaattaaacgaagtlttgagacgctt	1324		
Qy 781	cagcgcgtactcgtatcggcgaggtgtctgcacggttgcggttaacactgtctccacc	840		

Db	1325	ctggtgtttataacttaagtgtaagttltaacaaggaagatccaacttaatttlylcaatlc	1384
QY	841	agaacgtaatgagcgcgctactgtaactatccatccattactatccactccaagcgcttca	900
Db	1385	agaattatataagaagaggttacccaacttcatcatatbactataccagatataagattcttca	1444
QY	901	agtaaaccttcggcgcgcatgtgacgactcttcaacatgatcaaacaccgctcaaatccogact	960
Db	1445	gtgataacttcggcgacttcacglttcagtttaacttcaatgatctccagttacagttacatctt	1504
QY	961	gtccgacgttaacaacatccctgtggcagcatctgtcgagaacccagcaacccagcttgcgtt	10220
Db	1505	gttcgacgctcctcttgttgggaaccttattgaaacacatgcacgaagtatttccat	1564
QY	1021	cttaacccaagaacatagccctcgcgcgaagaagctgcagatcatcatcctcaagaagcg	1080
Db	1565	cgattactcagacacatcctcttgaattgaatlyacatlygcttlttaatttttggtygat	1624
QY	1081	gaatcccatcatcatctaagccgcgccaagaacagacactagccgcgggaagaacggccccgga	1140
Db	1625	gtatcccaattatttattatgtgccaagaagaagttcctaattgtgttccgcctcgcga	1684
QY	1141	accgcggaagcaacctggtcctcgggctaaccgcgcagacgagcgtlytaagaattaatg	1200
Db	1685	atagaagaagctttagtlytlaagtlygatataatccagttcagaatactacgagcctaaca	1744
QY	1201	cctcgcggaagcaacatccgggaacatgctccatgcgaagaatcagaattctcggcctca	1260
Db	1745	gtaaacataaatacaataagaataatcaagctcatltaagaagatcttcgcttcatccaacttca	1804
QY	1261	agaactggcccatctacaaaagcagacaacagatcgcctatgcgaagggagcgacagatggt	1320
Db	1805	aatccttcagttgttctctcttcagacaactataatagcactagaaagtgtagcgatg---	1861
QY	1321	cgcgagatgtygaactatttgtlccaacaaggtgccttcgsgtbatccgatacctctcct	1380
Db	1862	ctaataacatgattccattctttaaataatttagtttcaaaagggcaccagaatatbactg	1921
QY	1381	tgaatgtgtgcggtttacaaagccgcgcgaacatgaaacggaagttcatlygtctgcagcagc	1440
Db	1922	tcagaaacaacccggtctattctagtggtgataaagttatcatattattcttcgcaattccg	1981
QY	1441	tgaacggttggctcggatggagaatgtgcctgttcctatagtcgaaggtgsgtactaaggtat	1500
Db	1982	tttagctgtgtgactccggaagctctatcgtlatcaatttctggttggaatgcaccaagtt	2041
QY	1501	tgtatccgac	1510
Db	2042	acgctccgct	2051

RESULT	15
AAFl1273	
ID	AAFl1273 standard; cDNA; 605 BP.
XX	
AC	AAFl1273;
XX	
DT	13-MAR-2001 (first entry)
DE	Aspergillus niger EST SEQ ID NO:3796.
XX	
KM	Multiple gene expression; filamentous fungal cell; EST;
KM	expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KM	Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KM	culture condition; environmental stress; spore morphogenesis;
KM	metabolic pathway engineering; catabolic pathway engineering; ss.
XX	
OS	Aspergillus niger.
XX	
PN	WO200056762-A2.
XX	
PD	28-SEP-2000.
XX	

PF 22-MAR-2000; 2000MO-US07781.
XX
PR 22-MAR-1999; 99US-0273623.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MM, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
PI WPI; 2000-594572/56.
XX
DR
XX
PT Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XS
XS Claim 87; Page 1702-1703; 3161pp:English.

The present invention describes a method for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. The method uses fluorescence-labeled nucleic acids isolated from the FF cells and a substrate of expressed sequence tags (EST). The ESTs are used in the methods for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. Monitoring the global expression of genes from FF cells allows the production potential of the microorganisms to be improved. New genes may be discovered, possible functions of unknown open reading frames can be identified and gene copy number variation and stability can be monitored. The expression of genes can be used to study how FF cells adapt to changes in culture conditions, environmental stress, spore morphogenesis, recombination, metabolic or catabolic pathway engineering. Using ESTs provides several advantages over genomic or random cDNA clones including elimination of redundancy as one spot on an array equals one gene or open reading frame, and organisation of the microarrays based on function of the gene products to facilitate analysis of the results. AAF047478 to AAF11247 represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus niger*; AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are all specifically claimed in the present invention.

Sequence 605 BP; 146 A; 147 C; 156 G; 148 T; 8 other;

Query Match	21.3%	Score 368.8;	DB 21;	Length 605;
Best Local Similarity	97.28;	Pred. No. 3.5e-76;		
Matches 418; Conservative	0;	Mismatches 7;	Indels 5;	Gaps 4

[illegible]

QY 456 atggtcgatgtgtgtgtgctaaccatattggtctatgatgagccggtagctcagtcgattac 515
|||||
Db 359 atggtcgatgtgtgtgtgcta--ccatattgctatgatgagccggtagct- agtcgattac 415
|||||
QY 516 agtgtgttca 525
|||||
Db 416 agtgggttaa 425

Search completed: June 13, 2002, 17:29:12
Job time: 13493 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 13:45:42 ; Search time 3214.34 Seconds
(without alignments)
11288.978 Million cell updates/sec

Title: US-09-710-339-1
Perfect score: 1734
Sequence: 1 tcaatcaagctctccctc.....aaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
SUMMARIES					

1	1734	100.0	1734	6	AX140291	AX140291 Sequence
2	1689.6	97.4	1721	8	ASNAMY	D10461 A. shirouza
3	1083	62.5	3366	8	ASOAMY1	X12725 Aspergillus
4	1065.4	61.4	2337	6	ASNMTAA	D00434 Aspergillus
5	1063.8	61.3	2337	6	E02659	E02659 DNA sequenc
6	994	57.3	3160	8	AOAMY2	X12726 Aspergillus
7	992.4	57.2	3146	8	AOAMY3	X12727 Aspergillus
8	992.4	57.2	35244	8	AB021876	AB021876 Aspergill
9	989	57.0	2935	8	ASNMTAG1	M3318 A. oryzae Ta
10	985.8	56.9	2970	8	AF139925	AF139925 Aspergill
11	979	56.5	2443	8	ANMAYB	X52756 A. niger amy
12	908.2	52.4	2520	8	ANMAYA	X52755 A. niger amy
13	660	38.1	2016	23	E09025	E09025 cDNA encodi
14	434.8	25.1	2239	8	LK030376	U30376 Lipomyces K
15	409	23.6	2143	8	SCSMA2	X73497 S. cydessa
16	377.4	21.8	4215	6	E01174	E01174 DNA encodin
17	377.4	21.8	4223	8	SPALP1	X05791 Saccharomyc
18	377	21.7	1404	6	E03536	E03536 DNA sequenc
19	375.4	21.6	1404	6	E08057	E08057 DNA encodin
20	375.4	21.6	1404	6	E08060	E08060 DNA encodin
21	375.4	21.6	1404	6	I23833	I23833 Sequence 7
22	375.4	21.6	1404	6	I43344	I43344 Sequence 4
23	375.4	21.6	1404	6	I44693	I44693 Sequence 4
24	373.8	21.6	1404	6	E08058	E08058 DNA encodin
25	373.8	21.6	1404	6	E08059	E08059 DNA encodin
26	373.8	21.6	1404	6	I23830	I23830 Sequence 1
27	373.8	21.6	1404	6	I23831	I23831 Sequence 3
28	373.8	21.6	1404	6	I23832	I23832 Sequence 5
29	373.8	21.6	1404	6	I43341	I43341 Sequence 1
30	373.8	21.6	1404	6	I43342	I43342 Sequence 2
31	373.8	21.6	1404	6	I43343	I43343 Sequence 3
32	373.8	21.6	1404	6	I44690	I44690 Sequence 1
33	373.8	21.6	1404	6	I44691	I44691 Sequence 2
34	373.8	21.6	1404	6	I44692	I44692 Sequence 3
35	371.4	21.4	2160	8	S7586	S7586 Debarymyce
36	371.4	21.4	2161	8	SOAMY	X62079 S. occident
37	369.2	21.3	2293	8	SOAMY1	X16040 Schwannomy
38	364.8	21.0	3784	8	S38381	S38381 AMY-alpha-a
39	349.6	20.2	2290	6	A05233	A05233 Artificial
40	349.6	20.2	2290	6	A08557	A08557 Unknown alp
41	349.6	20.2	2280	6	BD000146	BD000146 Microorga
42	349.6	20.2	2280	6	E29394	E29394 Microorgan
43	315.8	18.2	9451	8	D83541	D83541 Cryptococcu
44	238.8	13.8	9451	8	AF208225	AF208225 Emerice11
45	223.8	12.9	2851	8	AB008370	AB008370 Aspergill

ALIGNMENTS

RESULT 1
AX140291 LOCUS AX140291 1734 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 1 from Patent WO0134784.
ACCESSION AX140291
VERSION AX140291.1 GI:14280549
KEYWORDS
SOURCE Aspergillus oryzae.
ORGANISM Aspergillus oryzae.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 1734)
AUTHORS blsg Rd-Frantzen,H., Svendsen,A. and Pedersen,S.
TITLE Fungamyl-like alpha-amylase variants
JOURNAL Patent: WO 0134784-A 1 17-MAY-2001;
Novozymes A/S (DK)
FEATURES
source location/Qualifiers
1..1734
/organism="Aspergillus oryzae"
/gb_xref="taxon:5062"
54..1550
/note="unnamed protein product"
/code_start=1

QY	1271	catctacaagaacgacacacacgcatcgccatgcgcaaggcgacagatgggtcgcatcgt	1330
Db	1261	CATCTACAAAGACGACACAGATGCCATCGCAAGGGCACAGATGGTGCAGATCGT	1320
QY	1331	gactatctgtgccacaagggtgcttgggtgattcgtataccctctccttgagtgtgc	1390
Db	1321	GACTATCTTGTCCAAAGGGGTGCGGTGGATTCGATACCCCTCCTTGAGTGGTGC	1380
QY	1391	gggttacacacgcccgcgcgcgcattgacgaaggatcattgctgcgcacgcgcgagttgc	1450
Db	1381	GGGTACACACACCGCCGACGCAATTGACGGAGGTCATTGGCTGCACGACCGTACGGTGG	1440
QY	1451	ttcgcgatgaaatgtgcctgtcttcctatgacgagtgaggctacctagggtattgcgac	1510
Db	1441	TTCCGATGGAATGATGCGCTGTTCCGATGCGAGTGGCGGTACCTAGGGATATTGATCCGAC	1500
QY	1511	tgagaaattggcaggttagcaagatcctgtagtagctcgtgaagggtggaagatatgatg	1570
Db	1501	TGAGAAAGTTGGCAGGTACCAAGATCTGTAGTACGCTCGTGAAGGGTGAGAGTATGTAGT	1560
QY	1571	gtactgctatcaatcgcgcgcgcattgacgaaggatgattgagttgattgacagttgagatcg	1630
Db	1561	GTACTGCTATTCATCTGGCATTTGGACAGTGAATTGAGTTTGATGTACAGTTGGAGTGG	1620
QY	1631	ttactgctgcatccctctatactcttcgattgttttcgaacctaatgccaagcagcgc	1690
Db	1621	TTACTGCTGTCATCCCTTATACTCTTTGATTGTTTTCGAACCTATATACCAAGCAGCAGC	1680
QY	1691	tagctctattataggaataaaaaa	1722
Db	1681	TAGCTATATATAGAAAGATCAGATTATATA	1712
RESULT	3		
AOAMY1			
LOCUS	AOAMY1	3366 bp	DNA linear PLN 10-APR-1993
DEFINITION	Aspergillus oryzae amy1 gene for alpha-amylase (BC 3.2.1.1).		
ACCESSION	X12725		
VERSION	X12725.1 GI:2430		
KEYWORDS	amy1 gene; amylase-alpha.		
SOURCE	Aspergillus oryzae.		
ORGANISM	Aspergillus oryzae		
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.		
REFERENCE	1 (bases 1 to 3366)		
AUTHORS	Wirsel, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-AUG-1988) Wirsel S., Inst. f. Biochemie, TH Darmstadt, Petersenstr. 22, D-6100 Darmstadt, FRG		
REFERENCE	2 (bases 1 to 3366)		
AUTHORS	Wirsel, S., Lachmund, A., Wildhardt, G. and Rutkowski, E.		
TITLE	Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon organization		
JOURNAL	Mol. Microbiol. 3 (1), 3-14 (1989)		
MEDLINE	89237897		
COMMENT	[2] also presents a cDNA seq (clone pGM8); pGM8 is lacking the first 42 bp of the exon when compared to genomic seq and includes a poly(A) tail.		
FEATURES	see X12725 - X12727 for A. oryzae amy1, amy2 and amy3 genes.		
source	Location/Qualifiers		
	1..3366		
	/organism="Aspergillus oryzae"		
	/strain="DSM63303"		
	/db_xref="taxon:5062"		
	/clone="lambda Stw7"		
exon	/clone_id="EMBL3"		
	552..782		
	/gene="amy1"		
	/number=1		
gene	552..2934		
	/gene="amy1"		
prim_transcript	552..2934		
	/gene="amy1"		
		sig_peptide	615..677
			/gene="amy1"
		CDS	join(615..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2658)
			/gene="amy1"
			/codon_start=1
			/product="alpha-amylase"
			/protein_id="CAA31218.1"
			/db_xref="GI:295921"
			/db_xref="SWISS-PROT:P10529"
		mat_peptide	/translation="MVAWMSLFYGLQVAAAPALAAPADMRSGITFLLDTRAPD GSTATCNTRADQKCGGTWGGIIDKLDTIGMGFTAIWTPVTAQLPQTPAYGDAYHG YWQDIDYSLNENGTADDLKALSSALHERGMYLAVDVVANHMGYDGAAGSSVDSEKPFSSODYPHPCFIONYEDQFOVEDCMGLNTVSLPDDTTRKDYKNEMYDMVGSLSN YSIDGLRIDFVKHVKDQFMPGNKAAGVYICIGVELDGDPAVTCQYKVMNDGVLNPTIY XPLLNKSTSGSMDLXNMINTYKSCPSPTILGTPVENHNDPRFASYNIDIALAKN VAAFTILNDISPIITLYAGQEHYAGNDPANREAWLSGFTPDSLYIKLISANIRNY AISKDNGEVYKMPYIKDDPTTLAMRGTDGSOIVTLLSNKGASGDSYTLISGAGYT AGQDLREIVICTVTVGSDGNVPMAGGLPRVLYPTEKLASGKICSSS"
			join(678..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2655)
			/gene="amy1"
			/product="alpha-amylase"
		intron	783..837
			/gene="amy1"
			/number=1
		exon	838..876
			/gene="amy1"
			/number=2
		intron	877..961
			/gene="amy1"
			/number=2
		exon	962..1077
			/gene="amy1"
			/number=3
		intron	1078..1146
			/gene="amy1"
			/number=3
		exon	1147..1255
			/gene="amy1"
			/number=4
		intron	1256..1323
			/gene="amy1"
			/number=4
		exon	1324..1552
			/gene="amy1"
			/number=5
		intron	1553..1610
			/gene="amy1"
			/number=5
		exon	1611..1773
			/gene="amy1"
			/number=6
		intron	1774..1838
			/gene="amy1"
			/number=6
		exon	1839..1985
			/gene="amy1"
			/number=7
		intron	1986..2050
			/gene="amy1"
			/number=7
		exon	2051..2291
			/gene="amy1"
			/number=8
		intron	2292..2370
			/gene="amy1"
			/number=8
		exon	2371..2934
			/gene="amy1"
			/number=9
		polyA_site	2934
			/gene="amy1"

BASE COUNT 924 a /note="pgm8" 865 c 755 g 822 t
ORIGIN
Query Match 62.5% Score 1083; DB 8; Length 3366;
Best Local Similarity 75.8% Pred. No. 1.5e-244;
Matches 1707; Conservative 0; Mismatches 0; Indels 544; Gaps 8;
QY 1 tcaatcaagctccctctctctgaacaataaaccccaagaagccttaataatgctg 60
DB 565 TCACATCAAGCTCTCCCTCTCTGAACATTAACCCACAGAAAGCATTTATGATGCTG 624
QY 61 cgtgtgtgtcctctatctctgtacagcctcagtcgagcagcctgtctgtcaacgc 120
DB 625 CGTGTGTGCTCTATTCTGTAGCGCCTTCAGTGCAGGACCGCTTGTGGCTCAACGC 684
QY 121 ctgagagctgagatgcgaatcattatctcctcgaagatcgattgcaagaagc 180
DB 685 CTGGGAGCTGGCGATCGCAATCCATTATTTCTTCAGGATGATTTGCAAGAGCGG 744
QY 181 atgggtcagacgtgagactgtatactgagatc----- 216
DB 745 ATGGGTGACGACTGCGACTGTGTAATCTGCGGATCAGGTGTGTTACCTACTAGCTT 804
QY 217 -----agaatactgtgtgagaatgcaagc 245
DB 805 TCAGAAAGAGGAATGTAACTGACTTGATATAGAAATACGTGTGTGGACATGCGAGGC 864
QY 246 atcatcgaca----- 255
DB 865 ATCATCGACAAGTAATATGGCCCTTATCAAAAAAGAGGAAAGCAGAAAGAAAA 924
QY 256 -----agttgacctatataccaaggaatggg 280
DB 925 TAAATTAAGAAAGAACTAGTCTTAACCATCAATAGTGTGACTATATCCAGGGAAAGGG 984
QY 281 ctcaagacacatcctgatacccccgltacagcccaagctgccccagacaacgcatatg 340
DB 985 CTTGACACACCACTGTGATGATACCCCGCTTACAGCCACAGTCCGCCAGACACCGCATATGG 1044
QY 341 agatgacctacatggtctactgagcaggaatat----- 373
DB 1045 AGATGCCCTTACCATGCTACTGGCAGCAGGATATTAAGTCGATTTCTTAAATATCTACC 1104
QY 374 -----atacctctgaaacgaaaa 391
DB 1105 TGTCACTCTTTACATCAATATGAACCTAACTTGATGTTTAGATACCTCTGAAGCAAAA 1164
QY 392 ctacgagacatgacagatgaaagcgctctcttcgccccctcatgagaagggagatgta 451
DB 1165 CTACGGCACTGCAGATGACTTGAAGGGCGCTCTCTCGGCCCTTATAGAGGGGAGTGA 1224
QY 452 tcttatggtcgatgtgtgtcttaacatat----- 481
DB 1225 TCTTATGTCGATGTGGTGTCTTAACCATATGTTCTGTGTCCTTTGCAACTGACTTCGCG 1284
QY 482 -----gggctatgtatgagggggtgagc 503
DB 1285 GATATGTTCAATTTCAGTACTGACAAATGATTAATATCAGGGGCTATGATGGAGCGGGTAGC 1344
QY 504 tcaagtcgaatcaagtggtttaaacggttcagttcccaagaactactccaccgcttctgt 563
DB 1345 TCAGTCGATTAAGTGTGTTTAAACCGTTTCAGTCCCAAGACTACTTCCACCGCTTCTGT 1404
QY 564 ttcattcaaaactatgaagatcgaactcaagttgaagatgtgcgtgtagagaataaact 623
DB 1405 TTCAATTAAGAACTATGAAGATCAGACTCAGCTGAGATTGCTGGCTAGAGATTAACACT 1464
QY 624 gtctcctgctgatactcgtatatacaaccaagaagatgtgtgcaagaatgatatgtaactgtg 683
DB 1465 GTCTCTTGCCCTGATCTGATACCAACCAAGAGATGTGTCAAGAAATGATGGTATGACTGG 1524

QY 684 gtggagatcatgtgtatcgaactactcca----- 711
DB 1525 GTGGATTCATTTGGTATCGAACTACTCGACTAGATATTCTTCCCTCATTTCAACTTGG 1584
QY 712 -----ttgagcgctccgtatcgcacaaagttaaacagct 745
DB 1585 CTGATGATGATACTTACGAAATCACTTATGACGGCCCTCCCTATCGACAGATTAACACAGT 1644
QY 746 ccaagaagacttcgccccgggttacacaagaacgcgaagcggtactatcgtgagaggt 805
DB 1645 CCAGAAAGACTTTCGCCCCGGGTACACAAAGCCGAGCGGTGTACTGTATGCGGAGGT 1704
QY 806 gctcgaagcttgatcgcgacctactgtlccclacagaagcgtcatgagcagcgactgaa 865
DB 1705 GCTTCGAGGGTATCCGGCTTACACTTGTCCCTACCAAGAGTATGAGAGCGGTACTGAA 1764
QY 866 ctatcccat----- 874
DB 1765 CTATCCCATGTATGTTCCTCCAAACCATGAGCCTTCTTGCAAGTCTCATCTCTAACGAA 1824
QY 875 -----ttactatccactcctcaagccttcaagttcaacctcgcggaagatg 920
DB 1825 ACGGCTTAAGAACAGTTACTATCCACTCTCTTAAGCCTTCAAGTCAACCTCCGGACAGATG 1884
QY 921 gacgacctctacacacatgaatcaaacacgltcaaatccgaactgtccagaactcaactctg 980
DB 1885 GACGACCTGTACACATGATCAACACCGTCAAAATCGAAGTGTCCAGACTCAACACTCCG 1944
QY 981 ggcacattcgtcgagagaacgaagcaaacccaggttcgctc----- 1021
DB 1945 GGCACATGTGTCGAGAAACGACGACCAACCCACGGTTCGTGTAAGTCTTCCCTTTAT 2004
QY 1022 -----ttacaccaagacga 1035
DB 2005 TTCCGTTCCCAATTTTCCACACAGAACCCCACTTACACAGCAAAAGTTACACCAAGCA 2064
QY 1036 taagcctcgcacaaagacgtcgagcatlcatcatcccaagaagatcccatcatct 1095
DB 2065 TAGCCCTCGCAAGAAAGTGCAGCATTTTCATCATCTCAACGACGGAATCCCATCATCT 2124
QY 1096 agcgcggtccaaagaacagcaactacgctgcggtggaagacccgcggaacgcggaagcaact 1155
DB 2125 ACCCGGCGCAAGAAAGCACTACGCGCGGGAACCAACCCCGCAACCCGCAAGCAACT 2184
QY 1156 ggtctcgggtcaccgcgacgagcagagctgttaagaagttaattgtcctcgcggaacga 1215
DB 2185 GGCTTCTGGGCTACCCGACCGACGACGAGCTGTACAGTTAATGTCTCGGAAAGGCA 2244
QY 1216 tccgaaactatgcatlaagcaagaatatacaggaattcgtgactaca----- 1260
DB 2245 TCGGAAACTATGCCATTAGCAAAAGATAGAGATGCTGACTACAAAGTAAACGACAACT 2304
QY 1261 ----- 1260
DB 2305 CTAAAGTACCCCTTAATGGCTATCTTCAAGATATCGACACAAAGACTTAATCACTGGCA 2364
QY 1261 -----agaactgtgccatctaaacaagaacgaacaagatgtgcattgtgcaaggacagat 1316
DB 2365 ATACAGAACTGGCCCATCTTCAAAAGACGACACACACATGCCATGCCCAAGGGCACAGAT 2424
QY 1317 gggctcgagatcgtgactactgtgtccaaacaaaggtgtctcgtgtgattcgtataccctc 1376
DB 2425 GGGTTCGAGATCTGTGACTATCTTGTCCAAAGAGGTCGTTCCGGGTGATTCGATACCCCT 2484
QY 1377 tcccttagtggtgcggtttacacagccggtccagcaattgaaggaagttcaattgtgtcagc 1436
DB 2485 TCTTTGAGTGTGCTGGGTATACACAGCCGGCAGCAATTTACGAGAGTGTCAATTTGCTGCACG 2544
QY 1437 accgttaaggttggttcogattggaatgtgctgttccatgtgcaaggtgtgactactaag 1496
DB 2545 ACCGTACGGTGTGCTCGGATGGAATGTGCTGTCTTATGGCAGGTGGCTACTTAAG 2604
QY 1497 gtaattatccgactgagaagttgtgcaggttagcaagaatctgtagtagctcgtgaagggtg 1556

```

|||||
Db 2605 GATATGATCCGACTGAGAGTGGCAGGTAGCATCTAGTAGCTCGTGAAGGGTG 2664
|||
QY 1557 gagagatatgtagtgcgtcattcaatcgtcgatggaagatgagttgattgatg 1616
|||||
Db 2665 GAGAGTAGTAGTAGGAGTACGTCTATTCATCTGCATGGACAGTGAAGTTGATG 2724
|||
QY 1617 tacagtggagtcgtctacgcgtcaccctatactcttcgattgatttcgaacct 1676
|||||
Db 2725 TACAGTTGAGTAGTGTACGTCTCATCCCTATACCTCTTGATTTTTCGAACCT 2784
|||
QY 1677 aatgccaagcagcagctagtctattataggaa 1707
|||||
Db 2785 AATGCCACAGCAGCTAGTCTATTATGAGAA 2815
|||||

RESULT 4
ASNTAA
LOCUS ASNTAA 2337 bp DNA linear PLN 01-FEB-2000
DEFINITION Aspergillus oryzae Taka-amylose A (EC 3.2.1.1),
alpha-1,4-glucan-4-glucanohydrolase[TA] gene, complete cds.
ACCESSION D00434
VERSION D00434.1 GI:217822
KEYWORDS Taka-amylose A; alpha-1,4-glucan-4-glucanohydrolase.
SOURCE Aspergillus oryzae (RIB40) genomic DNA, clone phi-AA107.
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
1 (bases 1 to 2337)
Tada, S., Jimura, T., Gomi, K., Takahashi, K., Hara, S. and Yoshizawa, K.
Cloning and nucleotide sequence of the genomic Taka-amylose A gene
of Aspergillus oryzae
Agric. Biol. Chem. 53, 593-599 (1989)
Compared with the amino acid sequence of mature TAA reported by
Toda et al., Trp-385 is inserted and Asp-476 is deleted.
FEATURES
Location/Qualifiers
1..2337
/organism="Aspergillus oryzae"
/db_xref="taxon:5062"
18..24
/feature="putative"
<118..285
/note="Taka-amylose A precursor (exon 1)"
join(118..285,341..379,465..580,658..766,835..1063,
1122..1284,1350..1496,1562..1802,1882..2169)
/note="Taka-amylose A precursor"
/codon_start=1
/protein_id="BA00336.1"
/db_xref="gi:217823"
/translation="MMVAMWSLFLYGLVAVAPALAAPADMRSSQSIYELLTRPARD
GSTATCMTADQKCGGTWOGIIDLKDIQMGFTAIWTPVLAQLPOTAYGAYHG
YWOODIYSLNENYGTADLKAISALHREGMYLAWDVANMGDAGSSVDYVFPF
ESSDYEHPREFIONYPEDQYOVCEWGLDNVSLDLDTPTDYVKNEYWVGSLSVN
YSIDGLRIDYKVKOKDFWPGYKNAAGVYCIGEVLDVDPATCYQVNMGVLPYLY
YPLNAFSTSGSMDLILNMINTWKSDCPDSTLGTVEBNDNRFASYNDDLAN
VAAFIILNDGPIIYAGCEQHYAGNDPANREAFMTGTPDSELYKLISANIRY
AISKDTGEVCTKNMPIYKDDFTIAMRGDGSQIVTILSNKAGSDSYTSLSAGIT
AGQQLTEVICTYTVGSDGNVPYPMAGLPRLVLPKLAGASICSSS"
121..180
sig_peptide
mat_peptide
join(181..285,341..379,465..580,658..766,835..1063,
1122..1284,1350..1496,1562..1802,1882..2166)
/product="Taka-amylose A mature peptide"
286..340
intrn
/note="Taka-amylose A intron 1"
326..331
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
341..379
exon
/note="Taka-amylose A precursor (exon 2)"
380..464
intrn
/note="Taka-amylose A intron 2"
441..447
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
465..580
exon

```

```

/note="Taka-amylose A precursor (exon 3)"
581..657
intrn
/note="Taka-amylose A intron 3"
639..644
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
658..766
exon
/note="Taka-amylose A precursor (exon 4)"
767..834
intrn
/note="Taka-amylose A intron 4"
814..819
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
835..1063
exon
/note="Taka-amylose A precursor (exon 5)"
1064..1121
intrn
/note="Taka-amylose A intron 5"
1095..1101
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1108..1113
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1122..1284
exon
/note="Taka-amylose A precursor (exon 6)"
1285..1349
intrn
/note="Taka-amylose A intron 6"
1331..1337
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1350..1496
exon
/note="Taka-amylose A precursor (exon 7)"
1497..1561
intrn
/note="Taka-amylose A intron 7"
1545..1551
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1562..1802
exon
/note="Taka-amylose A precursor (exon 8)"
1803..1881
intrn
/note="Taka-amylose A intron 8"
1862..1867
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1869..1874
misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1882..>2169
exon
/note="Taka-amylose A precursor (exon 9)"

BASE COUNT 630 a 598 c 535 g 574 t
ORIGIN
252 bp upstream of SalI site.

Query Match 61.4%; Score 1065.4; DB 8; Length 2337;
Best Local Similarity 75.3%; Pred. No. 2.1e-240;
Matches 1701; Conservative 0; Mismatches 6; Indels 552; Gaps 8;

```

[illegible][illegible]

TAX ADM AGENCY, JOZO SHIGEN KENKUSHO: KK
 OS Aspergillus oryzae
 PN JP 1990268685-A/1
 PD 02-NOV-1990
 PF 07-APR-1989 JP 1989086787
 PI TAKAHASHI KOJIRO, IIMURA MINORU, GOMI KATSUYA, HARA MASAMICHI,
 YOSHIZAMA KIYOSHI, TADA SETSUZO, TAMURA GAKUZO PC
 C12N15/56, C12G3/02, C12G3/03, C12N1/15, C12N9/30, (C12N15/56, PC
 C12R1.69),
 PC (C12N1/15, C12R1.69), (C12N9/30, C12N1.69):
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC *Source: strain=RI840:
 FH key Location/Qualifiers
 FT
 FEATURES
 source 1<..<2337
 1. 2337 /Product='takaamylase A',
 /organism='Aspergillus oryzae',
 /db_xref='taxon:5062'
 BASE COUNT 630 a 599 c 534 g 574 t
 ORIGIN

Query Match 61.3%; Score 1063.8; DB 6; Length 2337;
 Best Local Similarity 75.3%; Pred. No. 4.9e-240;
 Matches 1700; Conservative 0; Mismatches 7; Indels 552; Gaps 8;

QY 1 taaactcaagctctccctctctgacacaataaaccacagagacattatgctgcg 60
 Db 68 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAGCATTTTGAAGTCG 127
 QY 61 cgtggtgtctctattctctgtaagccttcagctgcgcgcactgcttggctgcaagc 120
 Db 128 CGTGGTGTCTATTTCGTACGGCTTCAGGTGCGGCACTCTTGGCTGCAAGCG 187
 QY 121 ctggcgactgcgcgtacgcacatcttattctctctcagatgcatttcagaagcgg 180
 Db 188 CTGGCGACTGGCGATGCGATCCATTATTCTCTCAAGGATGATTTGCAAGGACGG 247
 QY 181 atgggtcgaagcactgcagctgtaactgcgagc----- 216
 Db 248 ATGGGTGCGAGCTGCGATGCTTATTAATGCGGATCAGGTGCTTACTACTAGCTT 307
 QY 217 -----agaaatactgtgtggaacatgycagcgc 245
 Db 308 TCAGAAAGAGATGTAAACTGACTTGATATGAATACTGTGTGGAACATGGCAGGCG 367
 QY 246 atcatcgaca----- 255
 Db 368 ATCATGACAGAGTAATTGCCCTTTATCAAAAAAGAAAGAAAGCAGAGAAAAA 427
 QY 256 -----agttgacatataccagggaaatgg 280
 Db 428 TAAATATAAAGAACTCTGTCTCAACATCAATAGTTGGACATFATCCAGGGAATGG 487
 QY 281 ctccacagcaatctggatcaaccccggttaacagccagctgccccagacaacgcatatgg 340
 Db 488 CTTCACAGCCATCTGGATACCCCGTTACAGCGTGCAGCCAGCACGCAATATGG 547
 QY 341 agatgcctacatgctactatggcaagaaat----- 373
 Db 548 AGATGCTTACCATGGCTACTGGCAGCAGATATGTAAAGCATGTATTTCCTTTGAATA 607
 QY 374 -----atactcttg 383
 Db 608 TTAACCTGTATCTATTCTACATCATATGAACATACTGATGTTTATAGATACCTCTCG 667
 QY 384 aacgaactaagcgaatgcagatgaacttgaaagcgctctctcgccctcaatgaaggg 443
 Db 668 AACGAATACTAGCGCACTGCGATGACTTGAAGGCGCTCTCTTCCGCCCTTCATAGAGG 727

QY 444 gggatlatcttatgctcgatgtgtgtgtaaccatat----- 481
 Db 728 GGGATGATCTTATGGTCGATGGTGGCTTAACCATATATGTCGTGCTTTGCAACG 787
 QY 482 -----gggcatatggag 495
 Db 788 ACTTCGGGATGATGTTTCATTTCAGTACTGACAAATGATGATATATACAGGCGCTATGATGAG 847
 QY 496 cgggtgctcagtcgattcaatgctgtttaaaccgttcagttcccaagcactctccacc 555
 Db 848 CGGATGCTCAGTCGATTCAGATGATGATTTAAACGTTGAGTCCCAAGACTCTCCAC 907
 QY 556 cgttcgcttcatcacaacatgaagatcaagcactgaagttgaagatctgtcgaagag 615
 Db 908 CGTTCTGTTTCATTCAAAACTATGAAGATCAAGCTAGGTTGAGTATTTGCTGCTAAGAG 967
 QY 616 ataacactgtctctctgctcgtatcgtataccacaaagatgtgtcaagaatgaatggt 675
 Db 968 ATAACTGCTCTCTGCTGCTGATCTCGATACCAAGGATGTCAGAAATGATGCTG 1027
 QY 676 acgactggtgtggatcaatgtgtatcgaactactca----- 711
 Db 1028 ACGACTGGGTGGGATGATGTTGATCGAATCTCAGTAAGATATTTCCCTCATTTCTA 1087
 QY 712 -----ttgacggctccgtatcgacacagta 737
 Db 1088 CAACTTGGCTGATCGATGATGATCTTACGAATAATCGTTGACGGCTCCGTTATCGACACAGTA 1147
 QY 738 aaacacgtccagaagagactctgtcccggttacaacaagcgcgagcgtgtaactgtatc 797
 Db 1148 AAACAGCTCCAGAAAGGACTCTGCGCGGATACAAAGCCGCAAGCGCTGATCTATTC 1207
 QY 798 ggcgaggtgtctgacggtgtatcgcgcttaactgtccctaccagaagcgtatgagcgcg 857
 Db 1208 GCGAGGTGTGTGACGCTGTGTCGCGCTACACTTGTCCACGAAGCTCATGAGCGCG 1267
 QY 858 gtactgaactcccat----- 874
 Db 1268 GTACTGAACATATCCATGATGTTGCTCCCAACCATGAGCGCTTTCGAAGTCTCATTC 1327
 QY 875 -----ttactatcactcctcaagcctcaagtaactccg 912
 Db 1328 CTAAGAAAGCGCTAAACACAGTTACTATCACTCTCCCAAGCGCTTCAAGTCAACCTCCG 1387
 QY 913 gcagcatggaagcacttacaacatgatacaacgcgtcaatccagctgtccagactcaa 972
 Db 1388 GCAGCATGGAAGCCTCTTACACATGATCAACACCGTCAATCGATGTCAGACTCAA 1447
 QY 973 cactcctgggcacatcgtcgaagaacacgaacacacagctgcctc----- 1021
 Db 1448 CACTCCTGGGCACATTGCTGGAAGACCGACCAACCCAGCGTTGCTGTGATGCTTCC 1507
 QY 1022 -----ttaacg 1027
 Db 1508 CTTTATTTTCTGCTCCCAATTTCCACACAGAACCCCACTTAACAAGACAAAGTTACAC 1567
 QY 1028 caagacatagccctgcgcaagaacgtcgagacattcatcatcctaagaagaaatccc 1087
 Db 1568 CAAGGATATAGCCCTCGCAGAGAGCTGCGCATATCATATCTTACAGCAGGAATCCC 1627
 QY 1088 catcatctacgcgcgcaagaacagcactacgcgcgaggaagaaagaccccgcgaaaccgga 1147
 Db 1628 CATCATCTAGCGCGCCCAAGACAGCATGACGCGCGGGGGAAGACCCCGCAACCGCGA 1687
 QY 1148 agcaacctgctctcgggtacccgacgcagcagcgtgtacaaagttaattgtccctcgcg 1207
 Db 1688 AGCAACCTGGCTCCGCGTACCCGACCGACGAGACTGTATCAAGTTAATTCCCTCCGC 1747
 QY 1208 gaagcaatccggaactatgcattagcaagaatataagatttgatcactaca----- 1260
 Db 1748 GAAGCAATCCGGAATGATGCAATTAAGCAAGATACAGGATGCTGTGACTTCAAGGTAAAG 1807

OY	1261	-----	1260	exon	552..782	/gene="amy2"
DB	1808	CACAACCTTAAGCATACCCCTAATGCGCTATCTTCAGATATCTGCACAAAGAGACTAAT	1867	gene	552..2658	/number=1
OY	1261	-----	1308		/gene="amy2"	
DB	1868	CACGTGGCATACAGAACTGGCCCATCTTCAAAAGACAGACACACGATCCCATGGCCAGG	1927	CDS	join(615..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2658)	/gene="amy2"
OY	1309	gcacagatgggtgcagatcgtgtaactctgtccacaagggtgcttcgggtgattcgt	1368		/product="alpha-pre-amy2"	
DB	1928	GCACAGATGGGTCGAGATCGTACTATCTGTCCAAACAGGTCGCTCGGGTGATTCTG	1987		/codon_start=1	
OY	1369	ataccctctccttggtgtgtgcgggttacacagccgcgcgaattgaaggatcattg	1428		/protein_id="CAA31219.1"	
DB	1988	ATACCTCTCTCTTGAGTGTGGGTTCACAGCGCGCACCAATTGACGAGGTCATTG	2047		/db_xref="GI:1228077"	
OY	1429	gctgcacacccgtgacggttgcttcgagtggaatgtgcttcctcctaaggcaggtgagc	1488		/db_xref="SWISS-PROT:P10529"	
DB	2048	GCTGCACACCCGTGACGGTGGGTTCGATGGAATGTGCTTCTCTATGGCAGGTGGGC	2107		/translation="MMVAMWSLFLYGLQVAAPALATPADMRQSIVFLLTDREARTD	
OY	1489	taactaagggtatgtatccgaactggaagttgacagtagcaaatctgtagtgcgt	1548		GSTATACNTADQKRCYGGWOGIIDKLDYIOGCGFPAIWTPTTAOLPQTAAYGAYHG	
DB	2108	TACTAGAGGATATTGATCCGACTGAGAGTTGGCAGTAGAAGATCTGTAGTAGCTCGT	2167		WMOODIYSLNENYGTADDLKALSALSHENGYLAWDVVANHHGYGAGASSVAYSEFK	
OY	1549	gaagggtggaagatataatgagtactgtcttaacttgagcattggaagtgattga	1608	sig_peptide	FSOOEFHPEFCYIOWVEOETOVEDCMIGDNYSIDPLDPTKDYKNEBQYMGVSLYSN	
DB	2168	GAAGGGGTGAAGATATGATGATGTACTCTATTCATATTCGCAATTGGCAATTGGACAGTTTGA	2227		YSIDGIRITDYKHYOKDRPGRNKAAGYCIIEVLDDGPATTCPIQNMADGVLTNPIT	
OY	1609	gttataatgaagttgagtcgtactgtctgcatcccttactctcttgattgtttt	1668		YPLLNFKSTSSMDLLYMIINTVRSDEPDSLLTGFVENHNDNPRFATYNDIALAKN	
DB	2228	GTTTGAATGTACAGTTGGAGTGGTACTCTCTCATCCCTTATCTCTTCGATTGTTTTT	2287		VAAFIILNDGIPRIYAGQEQHYAGNDPANREATWLSQPTDSLEKLIASANAIRNY	
OY	1669	cgaacccaatgccaagcagcagctgctatatataggaaa	1707		AISKDTGEVTVKKNMPLIKDDPTIIMARKGDSQIVTILSNKASGDSYTLISLGAGYT	
DB	2288	CGAACCTTAATGCCAAGCACGCTAGTCTATTATAGAAA	2326		AGQOLTEVTCCTTVVGSQNVPEPMAGGLPRVLVPTFKLAGSKICSSS"	
RESULT	6				615..677	
LOCUS	AOAMY2		3160 bp		/gene="amy2"	
DEFINITION	Aspergillus oryzae amy2 gene for alpha-amy1ase (EC 3.2.1.1).				join(678..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2655)	
ACCESSION	X12726				/gene="amy2"	
VERSION	X12726.1	GI:2442			/product="alpha-amy1ase"	
KEYWORDS	amy2 gene; amylase-alpha.				783..837	
SOURCE	Aspergillus oryzae.				/gene="amy2"	
ORGANISM	Aspergillus oryzae				/number=1	
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				838..876	
REFERENCE	Wirsal,S.				/gene="amy2"	
AUTHORS	Direct Submission				/number=2	
TITLE	Submitted (26-AUG-1988) Wirsal S., Inst. f. Biochemie, TH Darmstadt, Petersenstr. 22, D-6100 Darmstadt, FRG				962..1077	
REFERENCE	2 (bases 1 to 3160)				/gene="amy2"	
AUTHORS	Wirsal,S., Iachmund,A., Wildhardt,G. and Rutkowski,E.				/number=3	
TITLE	Three alpha-amy1ase genes of Aspergillus oryzae exhibit identical intron-exon organization				1078..1146	
JOURNAL	Mol. Microbiol. 3 (1), 3-14 (1989)				/gene="amy2"	
MEDLINE	89237897				/number=5	
COMMENT	see X12725 - X12727 for A. oryzae amy1, amy2 and amy3 genes. No amy2 transcripts were detected in Northern blot analyses; exons and introns in RT are given accordingly amy1 gene (X12725) Data kindly reviewed (28-Mar-1989) by Wirsal S.				1147..1255	
FEATURES					/gene="amy2"	
source					/number=4	
	1..3160				/gene="amy2"	
	/organism="Aspergillus oryzae"				/number=6	
	/strain="DSM63303"				1324..1552	
	/db_xref="taxon:5062"				/gene="amy2"	
	/clone="lambda Stw26."				/number=5	
	/clone_id="EMBL3"				1553..1610	
	join(552..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2658)				/gene="amy2"	
	/gene="amy2"				/number=7	
					1611..1773	
					/gene="amy2"	
					/number=6	
					1774..1838	
					/gene="amy2"	
					/number=8	
					1839..1985	
					/gene="amy2"	
					/number=7	
					1986..2050	
					/gene="amy2"	
					/number=7	
					2051..2291	
					/gene="amy2"	
					/number=8	
					2292..2370	
					/gene="amy2"	
					/number=8	

exon 2371..>2658
 /gene="amy2"
 /number=9
 BASE COUNT 842 a 815 c 737 g 766 t
 ORIGIN
 Query Match 57.3% Score 994; DB 8; Length 3160;
 Best Local Similarity 74.8%; Pred. No. 1.4e-223;
 Matches 1618; Conservative 0; Mismatches 0; Indels 544; Gaps 8;
 QY 1 tcaatcaagctctccctctctctgaacaataaaccacagaagcatltaatgaatgctg 60
 DB 565 TCACATCAAGCTCTCCCTCTCTGTAACAAATAACCCACAGAGGCAATTTATGATGCTG 624
 QY 61 cgtggtgctctatcttcttgaaggccttcagtcgagcccttcgttgcagcagc 120
 DB 625 CGTGGTGGTCTCTATTTGTGACGGCTTCAGGTCGCGGCACTGCTTGGCTCAACGC 684
 QY 121 ctgagcagctgagcagcgaatccatcttctctcagcagcagcagcagcagc 180
 DB 685 CTGCGGACTGGCGATCGCAATCCATTTATTTCTTCTCAGGATCGATTTCAGAGACGG 744
 QY 181 atgggtcagcagctgagcagctgtaatactgcgagtc----- 216
 DB 745 ATGGGTGACGACGTCGACCTTGTAATACGCGGATCAGGTGCTGTTACTACTAGCTT 804
 QY 217 -----agaaatctgtgtgtgagaaatgagcagcagc 245
 DB 805 TCAGAGAAGAGAGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 864
 QY 246 atcatcgaca----- 255
 DB 865 ATCATCGACAAGGTAATTTGCCCTTTATCAAAAAAGAGAGAGAGAGAGAGAGAGAG 924
 QY 256 -----agttgactatatccaggaatgagc 280
 DB 925 TAAAAATAAAGAACTGATGCTTACCATCATCATATGATGATGATGATGATGATGATG 984
 QY 281 cttaacagcagcagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 340
 DB 985 CTTCACAGCCTGTGGATGACCCCGTTACAGCCCGTCCGACAGACCCGCAATATG 1044
 QY 341 agatgcctaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 373
 DB 1045 AATGCTTACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
 QY 374 -----atactctcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 391
 DB 1105 TGTCACTCTTTTACATCAATATGAACTGATGATGATGATGATGATGATGATGATGATG 1164
 QY 392 ctaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 451
 DB 1165 CTACGGCACTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1224
 QY 452 tcttaatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 481
 DB 1225 TCTTATGGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1284
 QY 482 -----gggcatatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 503
 DB 1285 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1344
 QY 504 tcaatgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 563
 DB 1345 TCAAGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1404
 QY 564 ttcattcaaaactgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 623
 DB 1405 TTCATTCAAAACTATGAAGATCAGACTCAGGTTGAGATGCTGCTGCTGCTGCTGCTGCTG 1464
 QY 624 gtctcctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 683

DB 1465 GTCCTCCCTGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1524
 QY 684 gtggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 711
 DB 1525 GTGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1584
 QY 712 -----tgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 745
 DB 1585 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1644
 QY 746 ccaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 805
 DB 1645 CCAAGAAGCACTTGTGGCCCGGTTACAAACCCGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1704
 QY 806 gctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 865
 DB 1705 GCTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1764
 QY 866 ctatcccat----- 874
 DB 1765 CTATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1824
 QY 875 -----tcaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 920
 DB 1825 ACGGCTAAACAGGTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1884
 QY 921 gacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 980
 DB 1885 GACGACCTCTACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1944
 QY 981 ggcacatcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1021
 DB 1945 GGCACATTCGTGAGAGAACACGACCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2004
 QY 1022 -----ttaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1035
 DB 2005 TTCCGTTCCCAATTTCCACACAGAACCCCACTTACAAAGAGAACTTACACCAACAGCA 2064
 QY 1036 taagc 1095
 DB 2065 TAGCCCTCGCGAGAAAGTGGAGATTCATCATCTTCAAGAGGAGATCCCATCATCT 2124
 QY 1096 aagc 1155
 DB 2125 ACCCGCGCAAGAACAGCACTACGCGCGGAAACACCCCGGAAACCCCGAAGCAACCT 2184
 QY 1156 ggcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1215
 DB 2185 GGCTCTCGGGCTACCCGACCGAGCGAGCTGTACAAAGTTAATTGCTCCGGAACGCA 2244
 QY 1216 tccggaactatgcatatgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260
 DB 2245 TCCGGAACCTATGCAATGAGAAAGATACAGGATTCGTGACCTTCAAGATGAGCAACACT 2304
 QY 1261 ----- 1260
 DB 2305 CTAAAGCATACCTTAATGCGCTATCTTACAGATGATGATGATGATGATGATGATGATGATG 2364
 QY 1261 -----agaactggccatctcaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1316
 DB 2365 ATACAGAACTGGCCCATCTTACAAAGAGACACACACATCTCCATGCGCAAGGCGACAGAT 2424
 QY 1317 ggcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1376
 DB 2425 GGGTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2484
 QY 1377 tccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1436
 DB 2485 TCCCTGAGTGGTGGGCTTACACAGCGCGCACCAATTTGACGAGGCTCATTTGGCTCCACAG 2544
 QY 1437 accgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1496
 DB 2545 ACCGTGACGGTGGTGGATGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2604

QY 1497 gttatgtatccgactgagaaattgacaggtagcaagatctgtactgtactgtgaagggtg 1556
|||||
Db 2605 GTATTGTATCCGACTGAGAGAGTTGGCAGGTAGCAAGATCTGTACTGTCTGTGAAGGtG 2664
QY 1557 gagaatcatgatgtactgtactatcaatctgcatctggaacagtgaattgattgatg 1616
|||||
Db 2665 GAGACTATATGATGTTACTGCTATTCATATCGGATTCGACAGTGGAGTTTGAGTTGATG 2724
QY 1617 ta 1618
||
Db 2725 TA 2726

RESULT 7
AOAMY3 3146 bp DNA linear PLN 10-APR-1993
LOCUS Aspergillus oryzae amy3 gene for alpha-amyase (BC 3.2.1.1).
DEFINITION X12727
ACCESSION X12727
VERSION X12727.1 GI:2454
KEYWORDS amy3 gene; amyase-alpha.
SOURCE Aspergillus oryzae.
ORGANISM Aspergillus oryzae.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
1 (bases 1 to 3146)
Wirsal,S.
Direct Submission
Submitted (26-AUG-1988) Wirsal S., Inst. f. Biochemie, TH
Darmstadt, Petersenstr. 22, D-6100 Darmstadt, FRG
2 (bases 1 to 3146)
Wirsal,S., Lachmund,A., Wildhardt,G. and Rutkowski,E.
Three alpha-amyase genes of Aspergillus oryzae exhibit identical
intron-exon organization
Mol. Microbiol. 3 (1), 3-14 (1989)
89237897
JOURNAL see x12725 - x12727 for A oryzae amy1, amy2 and amy3 genes. amy3
MEDLINE transcripts were detected in Northern blot analyses [2]. Data
COMMENT kindly reviewed (28-Mar-1989) by Wirsal S.
location/Qualifiers
1..3146
/organism="Aspergillus oryzae"
/strain="DSM63303"
/db_xref="taxon:5062"
/clone="lambda Stw28"
/clone_lib="EMBL3"
552..782
/number=1
614..2657
/gene="amy3"
614..676
/gene="amy3"
join(614..781,837..875,961..1076,1146..1254,1323..1551,
1610..1772,1838..1984,2050..2290,2370..2657)
/gene="amy3"
/codon_start=1
/product="alpha-amyase"
/protein_id="CAA31220.1"
/db_xref="GI:295922"
/db_xref="SWISS-PROT:P10529"
/translation="MNVAMWSLFYGLGVAAAPALATPADMRSOSIELLNDREFARTD
GSTTATCNFADKRYCGTMOGIIIDKLDTGCGFTAIWTPYTAOLPOTTAAGDAVHG
VMQODIYSINENYGTADDIKATLSALHENGWYLMVAVNHNHGYGAGSVDVSYEKP
ESSQDYFHFPCIONTERDOTQYEDCWLGNITYSLLDPTTKDYVKNERYDMVGSIVSN
YSIDGRIDTFKRVKQKDFWPGINKAAGVTCIGEVLDGPATYCPYQVMDGVLANPIY
YPLNAFKSTGSMDDLNMINTVKSDCDSTLIGTFVENHDNPRFASVYNDIALAKN
VAAFILINDGIPILYAGQSOHYAGNDPAMREATVLSGPTDSELYKLIASANAIRNY
AISKDFEFTVYKNNPIYKDDTTIARKKDGSGQIYVILSNKASGDSYTLISLSCAGYT
AGQGLEVIGCTTGYGSDGNVPMAGGLPRVILPTEKTLASIKSSS"
join(677..781,837..875,961..1076,1146..1254,1323..1551,
1610..1772,1838..1984,2050..2290,2370..2654)
/gene="amy3"
/product="alpha-amyase"
mat_peptide

intron 782..836
/gene="amy3"
/number=1
837..875
exon /gene="amy3"
/number=2
876..960
intron /gene="amy3"
/number=2
961..1076
exon /gene="amy3"
/number=3
1077..1145
intron /gene="amy3"
/number=3
1146..1254
exon /gene="amy3"
/number=4
1255..1322
intron /gene="amy3"
/number=4
1323..1551
exon /gene="amy3"
/number=5
1552..1609
intron /gene="amy3"
/number=5
1610..1772
exon /gene="amy3"
/number=6
1773..1837
intron /gene="amy3"
/number=6
1838..1984
exon /gene="amy3"
/number=7
1985..2049
intron /gene="amy3"
/number=7
2050..2290
exon /gene="amy3"
/number=8
2291..2369
intron /gene="amy3"
/number=8
2370..>2657
exon /gene="amy3"
/number=9
BASE COUNT 866 a 783 c 715 g 782 t
ORIGIN

Query Match 57.2%; Score 992.4; DB 8; Length 3146;
Best Local Similarity 74.8%; Pred. No. 3.4e-223;
Matches 1617; Conservative 0; Mismatches 1; Indels 544; Gaps 8;

QY 1 tcaactcaagctctccctctctgacacataaacccacagaagcatttgaatgctg 60
|||||
Db 564 TCACATCAAGCTCTCCCTCTCTGACATATAACCCGACAAAGCAATTATGATGTCG 623
QY 61 cgtgtgtctctatcttctgtacggccttcaggtcggacctgcttgcctcaagc 120
|||||
Db 624 CGTGTGCTCTATTCTGTACGCGCTTCAGTTCGCGCACCTGCTTGGCTGCACGC 683
QY 121 ctggcactggagatcggaatccatttcttcctctcagcagatcttgcaaggcgg 180
|||||
Db 684 CTGCGACTGGGATCGCAATCCATTATTTCTTCTCAGGATCGATTTCGAAGGACGG 743
QY 181 atgggtcagacactcgacttgtaatactgcggatc----- 216
|||||
Db 744 ATGGGTGACGACCTCGACTGTATTACTGCGGATCGGGTGTGTTACTACTAGCTT 803
QY 217 -----agaatactctgtgtggaacatgcaaggc 245

Db	804	TCAGAAAGAGGAAATGTAAACGACTTGATATAGAAAATACGTGTGGAAACATGCGAGGC	863
QY	246	atcatcgaca-----	255
Db	864	ATCATCGACAGAGTAAATGTGCCCTTTATCGAAAAAAGAAAGAAACAGACAGAAAA	923
QY	256	-----	280
Db	924	TAAATTAAGAAAGAACTAGTCTTACCATCACACTACTTGTGGACTATATCCAGGAAATGG	983
QY	281	cttcacagccatctygatcacccccgtttagagcccaagcttgcacccagacacacgcatalyg	340
Db	984	CTTTCACAGCCCATCTGGATGATCACCCCGCTTAAAGCCACAGCTGCCACAGACCCGATATGG	1043
QY	341	agatgcctcacatgagctgactgagcagggat-----	373
Db	1044	AATGCTCTACCATGGCTACTGTGGCAGCAGATATGTAAATGATGATTTCTTTAAATCTAC	1103
QY	374	-----	391
Db	1104	TGTCACTTTTACATCAATATGACACTAATGTAGTTGGTTTAAATATCTGTACAGAAAA	1169
QY	392	ctacagccatcgagatgacttgaaggcgctcttcgcgccctcalagaggggatg	451
Db	1164	CTACGGCACTCAATATGATTTGAAGGGCTCTCTTGGCCCTTCATGAGAGGGGATGTA	1223
QY	452	tcttatgctgagtgtgttgctaacatat-----	481
Db	1224	TCTATGTGATGTGGTGGTAAACCATATGGTTGCTGGTCTTTTGCAACTACTTCCG	1283
QY	482	-----	503
Db	1284	GATATGTTCACTTCACTACTGACAAATGATAATATACAGGCGCTATGATGAGCGGTAGC	1343
QY	504	tcaagtcgattacagtggtttaaacgcttcagttcccaagactacttccacccgcttct	563
Db	1344	TCAGTCGATTCAGCTGTGTTTAAACGTTTCCCAAGACTACTTCCACCCGTTCTGT	1403
QY	564	tctatcacaactatgaagtcagactcaggtttagagattgcttgcctaggaatacaact	623
Db	1404	CTCATTCAAAACCTTGAAGATTCAGACTCAGGTTGAAGATTTGCTGGCTAGAGATTAACCT	1463
QY	624	gtctccctgacctgcttcgataccacaaggatgtygtcaagaatgaatgtytaagactg	683
Db	1464	GTCCTCTTGCTGATCTCGATACACACAGATGTGTCAAGATGAATGGTACCACTGG	1523
QY	684	gtggagatcatgtytatcgaactactcca-----	711
Db	1524	GTTGGATCATTTGATGATGAACTCACTCCAGTAAGATATTTCTCCCTCATTTCAACAACCTTG	1583
QY	712	-----	745
Db	1584	CTGTATCGATGATTAATCGAATAAGTTGACAGGCTCTCCCTATCGACACAGTAAACACGT	1643
QY	746	ccagaagaagactcttgcgcccgggtacacaagaacgcgagcgtytacttlatcggcggagt	805
Db	1644	CCAGAAGAGACTTGTGGCCCGGCTGATCAACAACACCGAGCGGTATCTGTATCTGGCGAGGT	1703
QY	806	gtctgaagtgatccgggctcacttgttcctacagaagcgtcatlyagcggcgtactgaa	865
Db	1704	GCTGAGAGGTGATCCGCCCTACACTTGTCTTCAACAGAACGTCAATGAGCGCGTAACTGAA	1763
QY	866	ctaccccat-----	874
Db	1764	CTATCCATGTATGTTCCTCCACACCATATGACCTTCTTGGCAATGTCATCTCTCAACGAA	1823
QY	875	-----	920
Db	1824	ACGGCTAAACCAAGTTACTATCCACTCTCAACAGCCTTCAAGTCAACTCCGCGAGCATG	1883
QY	921	gagagactctaaacaatgatcaaacacgctcaaaatccgagctgtccagactcaagactcaactcgt	980

[illegible]


```

Db 21110 GATATGGTTCAATTCAAGTACTGACATGAGTAATATACAGGCGTATGATGAGCGGGTAGC 21169
Oy 504 tcagtcgattacagtggtgttaaacgcgttccagtcgttcccaagaactctccaccggttctgt 563
Db 21170 TCAGTCGATTCACATGGTGTAAACCGCTTCACAGTCCCAAGCACTACCTCCACCGCTTGT 21229
Oy 564 ttcatctaaactatgaagatcacagtcagtcaggttgtaggattctgtctgaagagataaact 623
Db 21230 CTCATTCAAAACCTATGAAAGATCAGACTCAGTGTAGAGATTCTGCTGATGAGATAAACT 21289
Oy 624 gtctctctgcgtatctcgatataccaaccaagaatggtgtcaagaatgaatggttaccgttg 683
Db 21290 GTCCTCTGCGCTGATCTGATACACCAAGGATGTGTCAGATGATGATGATGATGATGATG 21349
Oy 684 gtggagatcattggtatcgaaactatcca 711
Db 21350 GTGGGATCATTTGTATTCACACTACTCCAGTAATATTTCTCCCTCATTTACAACTTG 21409
Oy 712 -----ttgacgacctccgatacgacaagtaaacacgt 745
Db 21410 CTGATCGATGATCTTACGAATCAGTTGACGGCTCCGTTATCGACACACAGTAAACAGT 21469
Oy 746 ccaagaaggactctcgtcccggtgtacacaagaacgcgacgagcggtgactgtatcggcgaggt 805
Db 21470 CCAGAAGGACTCTTGCGCCGGGTACAAACAAGCCGACGCGTGTATGTATGTATGTATGTATG 21529
Oy 806 gctcgaagtgatccgagctcactgtccttaaccaagaagtcagtcagtcagtcagtcagtcag 865
Db 21530 GCTTCACGGTGTATCCGGCTTACCTTGTCTTACCAAGACGTATGAGCGGCGTACTGAA 21589
Oy 866 ctatcccat----- 874
Db 21590 CTATCCCATGTATGTTCTCTCCACACCATGAGAGCTTCTTGCAAGTCTCATCTCTCAAGAA 21649
Oy 875 -----tactatccactcctcaacgccttcaagtcacactccgcgaagcatg 920
Db 21650 ACGGCTAAAACCAAGTACTATCCATCCCTCAACGCGCTTCAAGTCAACCTCCGCGACGATG 21709
Oy 921 gacgagccttataacatgatacaaacgcgtcaaatccagacttccagactcaaacactcctg 980
Db 21710 GACGACCTCTTACACATGATGTCACACCGCTCAATCCGACTGTCTCAGACTCAACACTCCTG 21769
Oy 981 ggcacatcgtcgcagaaacacacacacacacacacacacacacacacacacacacacacac 1021
Db 21770 GGCACATTTGTGAGAGACACGACACCAACCGGTTCCGTTGTAGTCTTCCCTTTTAT 21829
Oy 1022 -----ttaaccaagacaa 1035
Db 21830 TTCCGTTCCCAATTTCCACACACAGAACCCACCTAACAGAGCAAAAGTTACACCAACGACA 21889
Oy 1036 tagcctcgcgaagaacgctgcgagcatcattcatctcaagaagcaatccccatcatct 1095
Db 21890 TAGCCCTCGCCCAAGACGTCGACATTCATCATCTCTCAAGACGGAATCCCATCATCT 21949
Oy 1096 acgcgcgcgaagaacacagcctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1155
Db 21950 ACGCGCGCCCAAGACACGACTACGCGCGGGAAGACGACCCCGCAACCGCGCAAGCAACT 22009
Oy 1156 ggcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1215
Db 22010 GGCTCTCGGGCTACCCGACGACGAGAGCTGTACAAAGTTAATGCTCCCGCAACGCAA 22069
Oy 1216 tcgcgacatacgcattagcaagaatagcagatggttgaacctaca----- 1260
Db 22070 TCCGGAATATGCCATTTACCAAGATACAGGATGTGTGACCTACAAAGGTAAAGCAACCT 22129
Oy 1261 ----- 1260
Db 22130 CTAAAGTACCTTAATGGGCTATCTTCAAGATATCGACACAAAGAGACTAATCACTGGGA 22189
Oy 1261 ----agaactggccactctacaagaacacacacacacacacacacacacacacacacacacac 1316
|||||

```

```

Db 22190 ATACAGACTGGCCCATCTACAAAGACGACACACAGATGCCATTCGCAAGGACAGAT 22249
Oy 1317 ggtcgcagatcgtgactatctgtcccaagaagggtgtcgcggtgattcgtatataccctc 1376
Db 22250 GGGTCCGACATGCTGACTATCTTGTCCAAAGAGGGGTGCTCGGGTATGATTCGATACCTC 22309
Oy 1377 tccctgaatggtggtgtgttacacagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1436
Db 22310 TCCCTGAGTGTGCGGGGTACACAGCCGCGCAGCAATTCACGAGGTATTTGGCTGCACG 22369
Oy 1437 accgtacggttggtgtcgcagatgaaatgtgtctgttccatataggcaggtggttaccctag 1496
Db 22370 ACCGTACGgttggtgttgcgatgaaatgtgtctgttccatataggcaggtggttaccctag 22429
Oy 1497 gtagtgcagctagagaatgtgcaggttagcaagatcgttagtgcgtgtgaagggtg 1556
Db 22430 GTATTTATCCGACCTGAGAAAGTTGGCAGGTAGCAAGATCTGTATGATCTGTGAAGGTTG 22489
Oy 1557 gagaatataatgattgattcgtatatacctgtgcatgtgacagatgattgtgattgtg 1616
Db 22490 GAGAGTATATGATGATCTGATTCATTCATCTGCAATGTCAGTGTGAGTTGATG 22549
Oy 1617 ta 1618
Db 22550 TA 22551

RESULT 9
ASNTRAG1 2935 bp DNA linear PLN 27-APR-1993
LOCUS ASNTRAG1 2935 bp DNA linear PLN 27-APR-1993
DEFINITION A.oryzae "Taka-amy1ase A (Taa-G1) gene, complete cds.
ACCESSION M33218.1 GI:166530
VERSION M33218.1 GI:166530
KEYWORDS Taka-amy1ase A.
SOURCE A.oryzae (strain JCM02239) DNA.
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurobiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 2935)
AUTHORS Tsukagoshi,N., Furukawa,M., Nagaba,H., Kirita,N., Tsuboi,A. and
Udaka,S.
TITLE Isolation of a cDNA encoding Aspergillus oryzae "Taka-amy1ase A:
EVIDENCE for multiple related genes
JOURNAL Gene 84, 319-327 (1989)
MEDLINE 90128276
FEATURES
source
location/Qualifiers
1..2935
/organism="Aspergillus oryzae"
/db_xref="taxon:5062"
/feature="Taa-G1 mRNA and introns"
543..>2789
/feature="Taa-G1 mRNA and introns"
607..669
/feature="Taka-amy1ase A signal peptide"
<607..772
/feature="Taka-amy1ase A (Taa-G1) precursor"
/number=1
/feature="Taka-amy1ase A (Taa-G1) precursor"
1603..1765,1830..1976,2041..2281,2360..2647
/feature="Taka-amy1ase A (Taa-G1) precursor"
/codon_start=1
/protein_id="AA032708.1"
/db_xref="GI:166531"
/translation="MMVAVWSLFLXGLQVAPALAAATPADMRSSQSYELLNDREARTD
GSTTATCNTADOKYCGGTWQGIIDKLDYIOGGMFTAIWTPVTAOLPOTTAAGDAVHG
WYQODIYSLNENYGTADDLKALSALHERGMYLMDVYVNHMGYDAGSSVYSEK
FSSQDFHPFCLIONEDQTOVEDCLGNVTYSLPDLTKDYVNEWYDVGSLVN
YSIDGLRIDTVKHVQDFMPGIRKAGVYCIQEVLDGDAATCPYONWMDVGLNPIY
YPLNLFKSTSGSHDLNININIVKSDCPDSTLLGTFVEMHNPFRASFTNDIAAKN
VAAFILINDGIPLIYAGOEHYAGNDPNNRATYASGYPTDSELYKLIASANAIRNY
AISKDTGFVYKMPYIKDDTTLAMRKGTGDSQIVTILSNKASGDSYTLISLGAQYT
AGQOLFVIGCTTVVSGDNVYPMVAGGLPRVLYPTEKLASGKICSSS"
join(670..772,828..829)
/product="Taka-amy1ase A"
mat_peptide

```

mat_peptide	join(772,828,.868)
intron	/product="Taka-amylose A"
exon	773..827 /note="Taa-G1 intron A (no splice consensus) ; putative"
intron	828..868 /number=2 /note="Taka-amylose A precursor"
mat_peptide	869..954 /note="Taa-G1 intron B"
exon	join(955..1070,1140) /product="Taka-amylose A"
mat_peptide	955..1070 /number=3 /note="Taka-amylose A precursor"
intron	join(1069..1070,1140..1248) /product="Taka-amylose A"
exon	1071..1139 /note="Taa-G1 intron C"
intron	1140..1248 /number=4 /note="Taka-amylose A precursor"
mat_peptide	1249..1316 /note="Taa-G1 intron D"
exon	join(1317..1545,1603..1604) /product="Taka-amylose A"
mat_peptide	1317..1545 /number=5 /note="Taka-amylose A precursor"
intron	join(1545,1603..1765,1830) /product="Taka-amylose A"
exon	1546..1602 /note="Taa-G1 intron E"
mat_peptide	1603..1765 /number=6 /note="Taka-amylose A precursor"
intron	join(1764..1765,1830..1976,2041) /product="Taka-amylose A"
exon	1766..1829 /note="Taa-G1 intron F"
mat_peptide	1830..1976 /number=7 /note="Taka-amylose A precursor"
intron	join(1975..1976,2041..2281) /product="Taka-amylose A"
exon	1977..2040 /note="Taa-G1 intron G"
mat_peptide	2041..2281 /number=8 /note="Taka-amylose A precursor"
intron	2282..2359 /note="Taa-G1 intron H"
exon	2360..>2647 /note="Taka-amylose A precursor"
mat_peptide	2360..2644 /product="Taka-amylose A"
polyA_signal	2784..2789 BASE COUNT 818 a 752 c 657 g 708 t ORIGIN

Query Match	57.0%;	Score 989;	DB 8;	Length 2935;
Best Local Similarity	74.7%;	Pred. No. 2.1e-222;		
Matches 1613; Conservative	0;	Mismatches 5;	Indels 541;	Gaps 8;

Accession	Sequence	Length
OY 1	tcacatcaagctctccctctctgaaacaataaacccacagaagcatattatgatgtcg	60
Db	TCACATCAAGCTCTCCCTCTCTGTAAACAATAAACCCACAGAAAGCATTTATATATGATGTCG	616
OY 557	ctgtgtgtctctatcttcgtacagccttaagtcgacgtgcttctgtcctgaacagc	120
Db	CGTGTGTCTCTATCTTCGTACAGCCTTAAAGTCGACGTCACCTGCTTTGCTGCACAGC	676

[illegible]

QY	712	-----	ttgaagcgcttcgtagtcgacacgaat	745
Db	1577	GCTGATCGATGATCTTACCAAAATCAGTTGACCGCTCCGTATCGACACAGTAAACACAGT	1636	
QY	746	ccagaagacattctggcccggtatacaaaacgcgacgagtcgtactatcgcggaagt	805	
Db	1637	CCAGAAAGACTTCTGACCCGGGTACAAACAAACCCGACGGGTACTATCGCGGAAGT	1696	
QY	806	gctcgacgctgtatcgcgacctacactgttcctcaccagaacgtatcatygcacgagctactaa	865	
Db	1697	GCTCGACGCTGATCCGGGCTTACACTTTGTCTCACCAGAAACGTACATGACAGCGGCTACTGAA	1756	
QY	866	ctatcccat-----	874	

Db	743	AAAAATAAATAAAGAACTAGTCTTAACCATTCACATAGTTGGATATATATCCAGGA				802
Qy	276	atgggttcacagccatctgtagatcaccccggttacagccagctgcccagaccgcga				335
Db	803	ATGGGCTTCACACCCATCTGGATCACCCCGTTTAAAGCCAGCTGGCCCAACACGCA				862
Qy	336	tatgagatgcttaccacatgagctacgtcgagagagatat				373
Db	863	TATGAGATGCTTCACATGGCTACTGCGACGAGGATATGTAGTCGATTTCTTAAATAT				922
Qy	374	-----				386
Db	923	CTACCTGTCACTTTTACATCAATATGAATGAATGATGGTTTATGATCTCTCTAAC				982
Qy	387	gaaaacragagcactcaaatgaatgtagagagcctcttcgcccctcatgaagagg				446
Db	983	GAAACATCAGGCACTGCAAGTGAAGTGAAGGCGCTCTCTGGCCCTTCATGAGAGGGG				1042
Qy	447	atgtatcttatggtcgaatggtgtgtgttacaatat				481
Db	1043	ATGTATCTTATGGTCATGGTGTGCTTAACCAATATGTTGCGGCGCTTTCACACTGACT				1102
Qy	482	-----				498
Db	1103	TCGCGATATGTTTCATTTCAGTACTGACAAATGAGTAAATATCAAGGGCTATATGATGAGCGG				1162
Qy	499	gtagctagtcgattacagtggtgtttaaacggttcagttcccaagcacttcacacct				558
Db	1163	GTACCTCAGTCGATTAACAGTGTTTTAAACCGTTCAGTTCCCAACATCTCCACCCGT				1222
Qy	559	tcgtttcatctaaactatgaatcagatcagatcaggttgaagattgtgcttagagata				618
Db	1223	TCTGTCTCATTTCAAAACATTAAGAAATCAACACTCAGGTGGAGATTGCGGCTAGAGATA				1282
Qy	619	acactgtctcttcgctcgatctcgatcagacacccaagagattgtgtcaagaatgaatgtagc				678
Db	1283	ACACTGTCTCTTBCCTCGATCTCGATACCAACAGAGATGTGTCAAGATGAATGTATCG				1342
Qy	679	actggtgtgagatcaatggtatcgaactactca				711
Db	1343	ACTGGGTGGGATCATTTGGTATCGAACTACTCCAGTAAATATTTCCCTCATTTCTCAAA				1402
Qy	712	-----				740
Db	1403	CTTGCGCTGATGATGATACTTACGAATAAGTTGAGGCGCTCGATGAGACAGTAAAA				1462
Qy	741	caagtcacagaagagactctgtgcccgggttacaacaaagccgcgaggtgttactgtatcgcg				800
Db	1463	CACGTCAGAAAGGACTTCTGTGGCCCGGATPAACAACAAAGCCGAGGGCTGTACTGATGGC				1522
Qy	801	gaagtgctcgaaggtgtatccgagctacactgttccctacagaacgctatggaagcggtat				860
Db	1523	GAGGTGCTCGACGGTGAATCGGGCTCACTGTCTCTTACCAAAAGCTATGAGCGCGTA				1582
Qy	861	ctgaacatcccat				874
Db	1583	CTGAACATATCCATGATGTTGTTCTCTCAACCATGAGCCTTTTGGAAAGTTCATCTCTA				1642
Qy	875	-----				915
Db	1643	ACGAAGCGGCTAAACACGATTACTATTCACCTCTCAACGCCCTTCAAGTCAACCTCCGCA				1702
Qy	916	gcattgagagaccttatacaatgatataaagctaaatcgcagctgttcagatctacaac				975
Db	1703	GCATTGAGAGACTCTTACAAATATATCAACCGTCAAAATCCAGCTGTCTCAGACTTAACAC				1762
Qy	976	tcctggtgacatctgttcgagagacacgacacccacaggtgtgcctc				1021
Db	1763	TCTGTGGCACTTGTGTGAGAACACAGACACCCACGTTTCCTTGTGTAAGTCTTCCTT				1822
Qy	1022	-----				1030

D81823	TTATTTTCCTCCCTCCCAATTTCCACACAGAACCCCATCTTAACAAGAGCAAGATTACACCA	1882
OY1031	cgcaatagcccttcgccaagaacgctgcgagcatctatcatctctaaagcgaatctcccat	1090
D81883	CGCATATGACCTTCGGCCAAACAGTGGCAGCATATCTCTCAACACAGGAATCCCAT	1942
OY1091	catctacgcgcgcgaagaacagactaacgcgcgcgcgaagaccccgcgagccgcgaac	1150
D81943	CATCTAGCCGGCGCAAGAACGACTACGCCGGGGGAAGACACCCCGCGAATCCGGCAAC	2002
OY1151	aaacctgcttcgggctactcccgaccgaagcgagctgtacaagttaatgtccctcgca	1210
D82003	AACCTGGCTTCGGCGTACCCGACGAGAGGTGTACAAAGTTAATTCCTCCCGCA	2062
OY1211	cgcaatccggaactatgacattagcaaaagatataagattctgtacata	1260
D82063	CGCATCCGGAATATGCCATTAGCAAAAGATACAGAGATTCTGTACATCAAGSTAAAGCAC	2122
OY1261	-----	1260
D82123	AACCTTAACGATACCCCTAATGCGCATATCTTACAGATATCTGACACAGAGACTAATAC	2182
OY1261	-----agaactggcccatctacaagaagacaacaagatcgcaatcgcaagggca	1311
D82183	TGGCAATACAGAACTGGCCCATCTTCAAAAGACGACACACATCGCATCGCAAGGGCA	2242
OY1312	caagatggctcgcagatcgtgactatcttgcacaagaagtgcttcgggtgattcgtata	1371
D82243	CAGATGGGTGCGAGATCGATGACTACTTGTCCAAACAGGGTGCTTCGGGTGATTCGTATA	2302
OY1372	ccctctcccttgatgtgtgcggtttacacagccgcgcgaactgtacgagaggtcattgct	1431
D82303	CCCTCTCTCTTGAATGGTGGGGTTPAACGCCGGCCACCAATTTAACGAGAGTATTGGCT	2362
OY1432	gcacgacccggaagtggttcgagatgagaatgtccgttcctcatatggaagtgctgac	1491
D82363	GCAGACCGGTGACGGTTGGTTCGAGATGGAATGTCCGTTCATATGGACAGTGGGGCTAC	2422
OY1492	ctaaggtatctatccgaactgaaagattggcaggttagcaagaatctgtatgactcgtgaa	1551
D82423	CTAAGGATATTGTATCCGACTGAGAAATTTGGCAGGTACGCAAGATCTGTATGATCGTGAA	2482
OY1552	gggtggagatataatgatgtactgtatctaatctggccatgtgacaggtgaattgatt	1611
D82483	GGGTGGAGATATATATGTACTGCTATTCATCTGGCATTTGGACAGTAGTTTGATT	2542
OY1612	tgatgta 1618	
D82543	TGATGTA 2549	
RESULT 11		
LOCUS	ANAMYB 2443 bp DNA linear	PLN 01-DEC-1992
DEFINITION	A.niger amyB gene for alpha-amyase.	
ACCESSION	X52756	
VERSION	X52756.1 GI:2325	
KEYWORDS	alpha-amyase.	
SOURCE	Aspergillus niger.	
ORGANISM	Aspergillus niger	
REFERENCE	Eukaryotes: Trichocomaceae; mitosporic Trichocomaceae; Aspergillus	
AUTHORS	Korman,D.R., Bayliss,F.T., Barnett,C.C., Carmona,C.L., Kodama,K.H.,	
TITLE	Royer,T.J., Thompson,S.A., Ward,M., Wilson,L.J., and Betka,R.M.	
JOURNAL	Cloning, characterization, and expression of two alpha-amyase	
MEDLINE	genes from Aspergillus niger var. awamori	
FEATURES	Curr. Genet. 17 (3), 203-212 (1990)	
SOURCE	Location/Qualifiers	
	1..2443	
	/organism="Aspergillus niger"	
	/strain="var.awamori, UVK143f"	

	BASE COUNT	669 a	620 c	558 g	595 t	1 others
polyA-site		2336	2341			
exon		/number=9				
		/gene="amyB"				
		1957 .	2244			
intron		/number=8				
		/gene="amyB"				
		1878 .	1956			
		/number=8				
		/gene="amyB"				
		1637 .	1877			
		/number=7				
		/gene="amyB"				
intron		1572 .	1636			
		/number=7				
		/gene="amyB"				

	BASE COUNT	669 a	620 c	558 g	595 t	1 others
polyA-site		2336	2341			
exon		/number=9				
		/gene="amyB"				
		1957 .	2244			
intron		/number=8				
		/gene="amyB"				
		1878 .	1956			
		/number=8				
		/gene="amyB"				
		1637 .	1877			
		/number=7				
		/gene="amyB"				
intron		1572 .	1636			
		/number=7				
		/gene="amyB"				

Db 871 GATATGCTTCATTTCAGTACGTCGAATGAGTATATCAGGGCTATGATGAGCGGGTACG 930
 Qy 504 tcaagtcgatacagatgctgtttaaacggttcagttcccaagactactccaccgcttcgt 563
 Db 931 TCAGTGTGATTACAGTGTGTTAAACCGTTCACTGCCAAGACTACTTCCACCCGTTCTGT 990
 Qy 564 ttcaatcaaaactatgaagatcaagactcaggttgaagattgcttggtcgaagataact 623
 Db 991 TTCATTCAAAACTATGAAATGACAGACTCAGTGTGAGGATTGCTGGCTAGAGATTAACCT 1050
 Qy 624 gtctcttcgctgatactgatacccaagatgtygttaaaagaatgtytaagactggt 683
 Db 1051 GTCTCTTGCTGCTGATCTCATTCACATCCCAAGAGATGTGTCAAGATGAATGTACAGCTGG 1110
 Qy 684 gtggagatcattggtatcgaactactca----- 711
 Db 1111 GTGGGATCATTTGGTATCGAATACCTCAGTAAGATTTTCTCCCTCATTTCAACTTGG 1170
 Qy 712 -----tgaagcgtccgtatcgatacagaataaacagt 745
 Db 1171 CTGATGATGATTAATTCAGAAATCAGTTCAGCGCTCCGCTATCGACACAGTAAACAGCT 1230
 Qy 746 ccagaagagactctgagccgggttaacaagaacgcgaagcgtgtactgtatcgagagct 805
 Db 1231 CCAGAAGGACTTCTGGCCCGGGTACACAAAGCCGAGCGTGTACTGTATGCGCAGGT 1290
 Qy 806 gctcgacggtgataccggtctactactgtccctacagaagcgtatgagaagcgtactgaa 865
 Db 1291 GCTCGAGGTGATTCGGGCTTACACTTGTCCCTACCAAGACGTCAATGGAGCGGTACTGAA 1350
 Qy 866 ctatcccat----- 874
 Db 1351 CTATCCCATGATGATGCTCTCCACACATGAGCGCTTCTGCAAGTCTCATCTCTTAACGA 1410
 Qy 875 -----ttaactacactcccaagcgttcaagtaactcaactcggcagactg 920
 Db 1411 AAGGCTAAACAGTACTTACATTCACCTCTCCACAGCCTTCAACTCAACCTCGGCGAGCTG 1470
 Qy 921 gaagactctacaatatgatacaacgcgttcaaatcgaactgtcgcagactcaaacactcgt 980
 Db 1471 GAGGACTCTTACACATGATGATCAACACGTCGAATCGAAGTGTCCAGACTCAACACTCTGT 1530
 Qy 981 ggcacattcgtctgagaacacagacacacacacggttcgtc----- 1021
 Db 1531 GGCACATTCGTGAGAACCCACGACCAACCCACGTTGCTTCTGTAAGTCTTCCCTTTATT 1590
 Qy 1022 -----taagccaaagaca 1035
 Db 1591 TTCGGTTCOCATTTCCACACAGAACCCACCTTAACAAGACCAAACTTACACCAACGACA 1650
 Qy 1036 tagcctcgcaagaacgttcgcagatcatcatcctcaacgacggaatcccatcatct 1095
 Db 1651 TAGCCCTCGCCCAAGAACGTCGACATTCATCATCTCTCAAGACGGAATCCCATCATCT 1710
 Qy 1096 agcgccgccaagaacagactcagcgcggaagaaagaccccgcgagacccgagaaacact 1155
 Db 1711 ACGCCGGCCAAACAGACTAGCGCGGGAAGAACGCCGCGAACCOCGGAACGCAACCT 1770
 Qy 1156 ggcctcgggtctaacccgagccgagcgtgtacaagttaatgtccctccgcgagacgaa 1215
 Db 1771 GGGCTCTGGGGCTACCCGACCGACGACGAGCTGTACAAATTATTTGCTCCCGGAACGCAA 1830
 Qy 1216 tccggaactatgcatagcaagaatacaggaatcgttactacta----- 1260
 Db 1831 TCCTGGAATATGCCATTAGCAAGATACAGGATTCGTGACCTACAAAGTAAAGCACAACT 1890
 Qy 1261 ----- 1260
 Db 1891 CTAAGCATACCTTAATGGCTTATCTTCAGAGTATCTGACACAGAGACTTAATCACTGSCA 1950
 Qy 1261 -----agaactggcccatctacaagaacagacaacagatcgcgaatgcgaagggcagagt 1316
 Db 1951 ATACGAACTGGCCCATCTACAAAGACGACACACAGATCCCGATGCGCAAGGGCACAGAT 2010

Qy 1317 ggcctcagatcgtgactatctgtccaagaagggtgcttcgggtgatactgtataccctc 1376
 Db 2011 GGGTCGACGATCGTACTATCTTGTCCAACAGGGGCTCTTGCGGTGATTCGTATACCTTC 2070
 Qy 1377 tccctgagtggtcgggtttacacagccgccaagaattgacgaggtgcatctggtcagc 1436
 Db 2071 TCCTTGATGTGGCGGTTACAGCGCGCCAGCAATTGACGAGGTCATTGGCTGTGACG 2130
 Qy 1437 accgtgacggttggttcggatggaatgtgcctgtccctatggcaggtyggactaagg 1496
 Db 2131 ACCGTGACGCTTGGTGGATGAGAAATGTGCTCTTCTTATGCGAGGTGGCTACAGG 2190
 Qy 1497 gtattgtaccgactggaagttgacgaggtagaagaatcgtgtagtctgttaagagtg 1556
 Db 2191 GTATTGATTCGGACTGGAATTTGGCAGGTACGACAAATCTGTATGACTGCTGTAAGGATG 2250
 Qy 1557 gagaatatagtgtagtctgtatcactcaactgtgcatgtgacgtgagttgagttgatg 1616
 Db 2251 GAGAGTATATGATGATGCTACTGCTATTCAATCTGGCATTGGACAGGAGATTAATGTGCTG 2310
 Qy 1617 tac 1619
 Db 2311 GAC 2313

RESULT 12

LOCUS ANMYA 2520 bp DNA linear PLN 01-DEC-1992
 DEFINITION A.niger amya gene for alpha-amyase.
 ACCESSION X52755
 VERSION X52755.1 GI:2323
 KEYWORDS alpha-amyase.
 SOURCE Aspergillus niger.
 ORGANISM Aspergillus niger.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

REFERENCE

1 (bases 1 to 2520)
 Korman,D.R., Bayliss,F.T., Barnett,C.C., Cazmona,C.L., Kodama,K.H.,
 Royer,T.J., Thompson,S.A., Ward,M., Wilson,L.J. and Berka,R.N.
 Cloning, characterization, and expression of two alpha-amyase
 genes from Aspergillus niger var. awamori
 Curr. Genet. 17 (3), 203-212 (1990)

JOURNAL

MEDLINE

90254827

FEATURES

Location/Qualifiers

1..2520
 /organism="Aspergillus niger"
 /strain="var. awamori, UVK143i"
 /db_xref="taxon:5061"

CAAAT_signal
 TATA_signal
 gene

exon
 /gene="amyA"
 201..368
 /number=1
 /gene="amyA"
 join(201..368,424..462,548..663,733..841,910..1138,
 1197..1359,1425..1571,1637..1877,1957..2241)
 /gene="amyA"
 /EC_number="3.2.1.1"
 /function="hydrolysis of alpha-1,4-glucosidic bonds in
 polyglucosans"
 /note="1, 4-alpha-D-glucan 4-glucanohydrolase"
 /codon_start=1
 /product="alpha-amyase-precursor"
 /protein_id="CA3696.1"
 /db_xref="GI:2324"
 /db_xref="SWISS-PROT:Q02905"
 /translacion="MMVAMWSLEFVIGLOVAAPALATPADMSOSIYFLINDREARTD
 GSTLATCNFADOKYCGTMOGILDKLDYIOGSGFTATITPTVNLQPTTAYGDAYHG
 YKQODYSINENYGRADDKALSSALHEGKRLMDVYANHHGTGAGSSVDYSYFKP
 FSSDYHFHCFIQTNETDQVEDEWLGNTVSLPDDTTKDQVVENENYDWGSLVSN
 YSIDGLRIDTVKHQKDFWPGYNKAGAYCIQSEVLIDBPAITCPCYQNMVDGLNPIY

sig_peptide
 mat_peptide
 /gene="amyA"
 join(264..368,424..462,548..663,733..841,910..1138,
 1197..1359,1425..1571,1637..1877,1957..2238)
 /gene="amyA"
 /EC_number="3.2.1.1"
 /note="1,4-alpha-D-glucan 4-glucanohydrolase"
 /function="hydrolysis of alpha-1,4-glucosidic bonds in
 polyglucosans"
 /product="alpha-amylose"
 369..423
 /gene="amyA"
 /number=1
 424..462
 /gene="amyA"
 /number=2
 463..547
 /gene="amyA"
 /number=2
 548..663
 /gene="amyA"
 /number=3
 664..732
 /gene="amyA"
 /number=3
 733..841
 /gene="amyA"
 /number=4
 842..909
 /gene="amyA"
 /number=4
 910..1138
 /gene="amyA"
 /number=5
 1139..1196
 /gene="amyA"
 /number=5
 1197..1359
 /gene="amyA"
 /number=6
 1360..1424
 /gene="amyA"
 /number=6
 1425..1571
 /gene="amyA"
 /number=7
 1572..1636
 /gene="amyA"
 /number=7
 1637..1877
 /gene="amyA"
 /number=8
 1878..1956
 /gene="amyA"
 /number=8
 1957..2241
 /gene="amyA"
 /number=9
 2242..2370
 2370..2375
 polyA_site
 BASE COUNT 682 a 645 c 581 g 611 t 1 others
 ORIGIN
 Query Match 52.4%; Score 908.2; DB 8; Length 2520;
 Best Local Similarity 73.6%; Pred. No. 2.3e-203;
 Matches 1537; Conservative 0; Mismatches 8; Indels 544; Gaps 8;
 QY 1 tcacatcaagctctccctctctgaacaataaaccacacagaagcattatgatgtcg 60
 |||

Db 151 TCACATCAAGCTCTCCCTCTCTGAAACAATAAACCCACAGAGGCATTATGATGTCG 210
 QY 61 cgtgtgtgtctctatattctgtacaggccttcaagtcgcgcgaactcttggctgcaagc 120
 |||
 Db 211 CGGTGTGTCCTATTTCTGTACGGCTTCAGTGTGGCAGCTCTTGGCTGAACGC 270
 QY 121 ctgcgactctgcagatcgcacatcatatttctctcagagatctgattgcaagcagc 180
 |||
 Db 271 CTGCGGACTGGGAGATCGCAATCATTTATTCCTCTCAGGATGATGTAAGAGCAGC 330
 QY 181 atgggtcgacgactgcgactgttaactcgcgac----- 216
 |||
 Db 331 ATGGGTGACGACTGGGACTGTGAATACGCGGATCGAGTGTGTTACTACTAGCTT 390
 QY 217 -----agaatactgtgtgtgaacatgtgcagggc 245
 |||
 Db 391 TCAGAAAGAGATGTAACCTGATTTGATGAATAATCTGTGGGAACATGCGCAGGGC 450
 QY 246 atcatcgaca----- 255
 |||
 Db 451 ATCATCGACAGGTAAATTGCCCTTTATCAAAAAAAGAGAAAGCAGAGAAAAA 510
 QY 256 -----agttgactatatccaggaatgg 280
 |||
 Db 511 TAAATTAAGAACTCTAGCTTAACATCAATCAATGTTGGACTATATCCAGGGAATGG 570
 QY 281 ctccacagccatctgtatcaccccgcttacagccagctgcgcccaagcaccgcgacatgg 340
 |||
 Db 571 CTTCACAGCATGTGATATCACCCCGTTACAGCCAGCTGCCAGACCGCATATGG 630
 QY 341 agatgcctacatgtgctactgtgcagcagatat----- 373
 |||
 Db 631 AGATGCTACCATGCTACTGCGCAGCAGATATGTAAGTCATTTCTTAATATCTACC 690
 QY 374 -----atactctgaacgaaa 391
 |||
 Db 691 TGTCAATCTTTTACATCAATATGACTAATCTTATGCTTTTAACTACTCTCTGACGAAA 750
 QY 392 ctacgcacatcagatcagatcagagcgtctcttcgcacctcatagagggagatgta 451
 |||
 Db 751 CTACGGGACTGCAGATGACTGTGAAGGCGCTCTCTGCGCTTCATGAGAGGGGATGTA 810
 QY 452 tcttaagtctgactgtgttgctaacatat----- 481
 |||
 Db 811 TCTTATGCTGATGTGTTGCTTAACCATATGTTGTTGCTTGTGCAACTGACTTCGCG 870
 QY 482 -----ggcctatatgagcggatgc 503
 |||
 Db 871 GATATGTTTATTTCACTACTGACAATGAGTAAATATCAGGCGCTATGATGAGCGGTACG 930
 QY 504 tcagtcgattacagtggtttaacggttcaagtcagttcccaagactactccaaccgtctgt 563
 |||
 Db 931 TCAGTCGATTACAGTGTGTTAAACCGTTCAGTCCCAAGACTACTCCACCGCTCTGT 990
 QY 564 ttcattcaaaactatgaagtcagactcaggttggaagattgtcgtggttagaagtaacact 623
 |||
 Db 991 TTCAATTCAAAACATGAAATGAAATCAGACTCAGCTTGAAGATTGCTGAGGAGTTAACCT 1050
 QY 624 gtctccttgctgactcgtatcaccaacaagagatgtgtcaagaatgaatgttgcagctgg 683
 |||
 Db 1051 GTCTCTTGCTGCTGATCTCGATACACCAAGATGTGTCAGAAATGAATGTACGACTGG 1110
 QY 684 gtggagatcattglatcgaactactca----- 711
 |||
 Db 1111 GTGGATCATTTGATGCACTACTCCAGTAAGATATTTCTCCCTCATTTCTACACTGG 1170
 QY 712 -----ttgacggctctcgtatcgaacagttaaacagct 745
 |||
 Db 1171 CTGATCGATGATACTTACGAATTCAGTTCAGCGGCTCTCCGATGACACATTAACACGT 1230
 QY 746 ccagaaggaactcttgcgcgggttacacaagaacgcagcagcgtgttactgttgcgcgaggt 805
 |||
 Db 1231 CCAGAGGACTTCTGGCCCGGTACAAACAGCCGCGCTGTACTGTTATGCGGCAAGGT 1290

[illegible]

XX	07-OCT-1997 (Rel. 52, Created)
DT	02-SEP-2000 (Rel. 65, Last updated, Version 2)
XX	
DE	cdna encoding acid-resistant alpha-amylase from Aspergillus kawachi.
XX	
KW	JP 1995099979-A/1.
XX	
OS	unidentified
CC	unclassified.
XX	
RN	[1]
RP	1-2016
RA	Oba T., Sudo S., Kaneko A., Tamura G.;
RT	"NMW GENE, TRANSFORMANT USING THE SAME AND USE THEREOF",
RL	Patent number JP1995099979-A/1, 18-APR-1995.
XX	TAX ADM AGENCY, JOZO SHIGEN KENKYUSHO:KK.
XX	
OS	Aspergillus kawachi
CC	JP 1995099979-A/1
CC	PD 18-APR-1995
CC	PF 30-SEP-1993 JP 1993265447
PI	Oba TOSHITERO, SUDO SHIGEOHSHI, KANEKO ANIHIRO, TAMURA GAKUZO
CC	PC C12N15/09,C12N1/19,C12N9/28,C12N15/09,C12R1:66),C12N1/19,
CC	C12R1:865),
CC	PC C12N9/28,C12R1:865);
CC	strandedness: Double;
CC	topology: Linear;
CC	hypothetical: No;
CC	anti-sense: No;
CC	Location/Qualifiers
CC	Key
CC	source
CC	1..2016
CC	/organism="Aspergillus kawachi"
CC	/strain="IFO 4308"
CC	1..2016
CC	/note="acid-resistant alpha-amylase from
CC	Aspergillus
CC	kawachi"
CC	Location/Qualifiers
CC	source
CC	1..2016
CC	/db_xref="taxon:32644"
CC	/organism="unidentified"
CC	Sequence 2016 BP; 470 A; 591 G; 493 G; 462 T; 0 other;

Query Match:	38.1%;	Score 660;	DB 23;	Length 2016;
Best Local Similarity	66.5%;	Pred. No. 6,96-145;		
Matches 945;	Conservative	0;	Mismatches 475;	Indels 0;
				Gaps 0;
QY	121	ctggagactgagcatcgcgaatccattatcttccttcacagatcgaattgcaagagcg	180	
Db	71	CTGAGAGATGGCCCATCTCAATCCATCTTCTTTCGCGGATCGGTTGGTAGACGG	130	
QY	181	atggatcagacagctgcgaattgaaactgacgacgaaatgactggtggaacatggc	240	
Db	131	ACAATTGCACTACAGTACGTGCATATACGGGAGACCAATCTACTGTGGTGAAGTTGGC	190	
QY	241	agggaatcatcgacaagtttgactatataccaggaatgggcttcaagccatctgataa	300	
Db	191	AAGGAATTATCAACCTCTGGACTATATCCAGGGGATGGATTACAGCATTTGGATCT	250	
QY	301	ccccccgttaacggccagctgccccagacaacagcatatggagatgacctaccatggtact	360	
Db	251	CGCCTATCACTAGAGACGTACCCAGGATCTTCGGATGGTGAAGCCTTACCATGGATACT	310	
QY	361	ggcagcagatatatactctctgaaacgaanaactaagcagctcagatgacttgaagcgc	420	
Db	311	GGCAGCGAAGATATATCTATGTGAATCTCCAACTTCGGACGCGAGATGATCTGAATGCC	370	

OY	421	tctcttcggcccttcacatgataaggggagtatactctatgtgtcatgtgtgttcataccata	480
Db	371	TCCTCCGATGCTCTTCACGGCCCGGGAAATGACTCATGTGTGACAGCTGCTCCCTACACCA	430
OY	481	tggagctatgtagagcggtgtagcttcagtcagttacagtgatgtatttaaacggttcagttccc	540
Db	431	TGGGCTACGGACGGTAACGGCAACGATGTGATTTACAGCTCTTTCGACCCCTTCGACTCCT	490
OY	541	aagactacttcaccacgcttctgttctatcttcaaaacatbtaagatcagactcagtttagg	600
Db	491	CCCTCCTACTCTCCATCCATCTACTGCTGCCTCATACAGATTGGGCAACTTGACCATGTGTCGAAG	550
OY	601	attctgtgtcagagataaacaactgtctcctgtccctgtatctgatacccaagaatgtg	660
Db	551	ACTGTTGGAGGGTGACACCATCATGTGCTCTCTCCCAATCTGAAACACACAGGAAACGGCG	610
OY	661	tcaagaatgaaatgtacagctggtgtgatcatctgttatcgcagactcactcatcagtcggcc	720
Db	611	TGAGAACCATTTGGTACGATTTGGGTAGCCGACTGGTATTCACACTACTAGTCGAGGGCC	670
OY	721	tcgclatcgacacaglaaaaaacgltcccaagaagactctgtgcccgggtlcaaaaacaagccg	780
Db	671	TCCGATGACACAGTGTGAGAAAGATGAAACCCGACTTCTCCGGGCTACCAAGAAGACAG	730
OY	781	cagcgctgtactgtatctcggcgaggtgtctcgaagtgatccggtcactcactgttccctacc	840
Db	731	CAGGAGTCTACTGTGGTGGTGGAATTCACAAACGGCAACCTGCTTCGACTCCCATACCC	790
OY	841	agaagctatcgaagcggtcactgtaacatcccaattacatccactccctcaagcgtcca	900
Db	791	AAAAATATCTAGATGGTGTCTCAACTATCCCATCTACTGCGAACACTCTCTACGGCTTTG	850
OY	901	agtcacactccggcgaagatctgagcactcctacacatlgatcaaacacgtlcaaatccagact	960
Db	851	AATCCTCCAGCGGCGACGATCGACAACTCTACAAATGATCAAAATCGTCGGCAGAGACT	910
OY	961	gtccagactcaaacactcctctgggacatctgtcgaagaaccaagcaaccgaagttcggctt	1020
Db	911	GCTCGATCCACACCTCTCTGGGCAACTTTATCGAAAAACACGCAACCCCGCTTGGCT	970
OY	1021	cttaacccaagacatagccctcccgcaagaagctcgagatcatcatccctcaacagag	1080
Db	971	CTGTACACATCCGACTACTCCCAACCCAAAACGTTCTCACTACATCTTCTCTCCGACG	1030
OY	1081	gaatcccatcatctacgcccggccaagaacagcactacgcccggcggaaaaaacaccccgca	1140
Db	1031	GCATCCCATGCTGTACGGCGGGCGGGAAGAACAGCACTACTCTCGGGCGGACAGTGCCCT	1090
OY	1141	accgcgaagcaacctgtgctctcgggtaccacccgacagcagcagctgtatacaagttaatg	1200
Db	1091	ACAAAGCGCATGTGCTGGCTATACAGGCTACACACTCTCCGGGGAGCTCTACACTGGATAG	1150
OY	1201	cctcgcgaagcgaatcccggaacatbcaacttagcaaaagataaagattcgtcgactaaca	1260
Db	1151	CCATCCAAAGCGCATCCGGAACTAGCTATTCTCAGCAGACTGGACTACATTTACTTTACA	1210
OY	1261	agaatgtgcccatctacaaagaacagacacacgatacgcactcgcgaagggcacaagatgagt	1320
Db	1211	AGAAAGACCCAATGTACACAGACAGCAACACATCGGGATGGCCAAAGGCACTCCGGCT	1270
OY	1321	cgcagatcgtgaatactctgtccaacaagggtcttcgggtgatactgataccctcctc	1380
Db	1271	CCCAATATCATCACGTCCTCTCCAAACAAAGGCTCTCCGGAAACACACTACACCTTCACCC	1330
OY	1381	tgaatgtgtgcggtttacaacgcccggccagcaatltgacggaaggtcatctgtgtcagacag	1440
Db	1331	TCAGGGAGGGCGGTACACGTCGCCGACGAAGCTCATCGAAGGCTACACTCTGACGTCGG	1390
OY	1441	tgaagctgtgttcggatgaaatgtgccttctcattgagcaggttggtgtactcaatgatat	1500
Db	1391	TGACGCTGACTCCAGACGGGATATCTCCTGTGCGGATGCGTTGGGGATTAACCTAGAGTTTC	1450
OY	1501	tgatccgactgagaadtgtggcaggttagcaagatctgtag	1540

Db	1451	TCCTCCCTGCTTCGGTGGTGTGATATGTTCTTCGCTTTGTGG	1490
RESULT 14			
LKU30376			
LOCUS	LKU30376	2239 bp	mRNA linear
DEFINITION	Lipomyces kononenkoae subsp. spenceremattinsiae alpha-amylose mRNA,		
ACCESSION	U30376		
VERSION	U30376.1	GI:1173536	
KEYWORDS			
SOURCE			
ORGANISM	Lipomyces spenceremattinsiae.		
REFERENCE	Lipomyces spenceremattinsiae		
AUTHORS	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
TITLE	Saccharomycetales; Lipomycetaceae; Lipomyces.		
JOURNAL	1 (sites)		
REFERENCE	Steyn, A.J., Marmur, J. and Pretorius, I.S.		
AUTHORS	Cloning, sequence analysis and expression in yeasts of a cDNA		
TITLE	containing a Lipomyces kononenkoae alpha-amylose-encoding gene		
JOURNAL	Gene 166 (1), 65-71 (1995)		
REFERENCE	2 (bases 1 to 2239)		
AUTHORS	Pretorius, I.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-JUN-1995) Isak S. Pretorius, Microbiology, University		
FEATURES	of Stellenbosch, Victoria, Stellenbosch, Western Cape 7600, South		
source	Africa		
	Location/Qualifiers		
	1..2239		
	/organism="Lipomyces spenceremattinsiae"		
	/strain="IC64052B"		
	/sub_species="spenceremattinsiae"		
	/db_xref="taxon:56873"		
	/lab_host="Saccharomyces cerevisiae"		
	242..1954		
	/codon_start=1		
	/product="alpha-amylose"		
	/protein_id="AAC49622.1"		
	/db_xref="GI:1173537"		
	/translation="MCGSTLSASLYVYNDYDKITVTLTYLTSSGTTGSTALLLPVMS		
	NNWELMTLSAIAAGAVETTGASVYDSVDSPTVYTLSDLPITTSASVPTGTRAMRGK		
	SIYOVDRPARKPTGSGTISCDVDVDRYCGSVSGYGIINMLDIQGMPTATVTSPIVE		
	NIPTDGTGIVAHGIMWKADITALENTREGADDLALATELHNRMKYLKMDLVVNHFAR		
	SGHNADVYSELPYPSQDYSQDFHCWITIDYNSQTNVEECWLGDSVPLVDVNTGLDTV		
	KSEYDQWVKOLIANYSIDGLRIDYKHMDFMFPQAFAGIYVGEVDFDPSYTCDE		
	YOENLDVLANPYVYVYAFORGVGSISLVDMIDTLEKSCIDPTLLGSLFLEMDNPE		
	RPESTYDSESLIKNAIAFTLISDGIPIIYCOBGLNGNDPNYRMLMPGYSFTTSS		
	FYEYIASLNLQNRNAIITYDPTLYTYQNNVYISDSFTTIAKRGFTGNQIITYLSMGSS		
	GSSTYTLSTNGTVASSVYVEILCTAVTVDLISGLAVPMGSLPRVHYPSQLVGSQ		
	ICSMV		
BASE COUNT	549 a 488 c 491 g 711 t		
ORIGIN			
Query Match	25.1%;	Score 434.8;	DB 8; Length 2239;
Best Local Similarity	56.2%;	Pred. No. 8.2e-92;	
Matches 817; Conservative	0;	Mismatches 637;	Indels 0; Gaps 0;
QY 97	cgggacacgcttcttgctgtgaacgacctgcgagatggtgcgataccatcatttcttc	156	
DB 504	CCGCGCTGTCCCTACAGGACACGCTGCAAAATGGGAGAGTGAAGCTATCATCAAGTTCG	563	
QY 157	tcacgactgatttcaagagacgagtggctgacgacgtcgactgttaatactgcgagtc	216	
DB 564	TAACTGATAGATTTCGACGCACTGATGCTGCATTTCATATTCATGCGACGTCACCGA	623	
QY 217	agaatactgttggtgaacatgacgagcagatcatcgacaagtgtgactatccagggaa	276	
DB 624	GGGTGATTTGGCGAGGGCTTTACCGGGGAGATCATCAACTGCTGGATTTCATCCAAGCA	683	
QY 277	tgggtctcaagcgcattctgatacccccgcgttaacagccagctgcgccacacgcgcat	336	

D	b	684	TGGGTTTAACTGCTATTTTGGATATCTCCCTATAGTGGAAATATTCGGATACACCGGCT	743
Q	y	337	atgagatgcttaccatcctacgtgctacgtgcagcaggaatataactctctgaaacaaactacg	336
D	b	744	ACGGTTACGGATATCATTGGTTATTGGATTGAAGATATCTTCGCCCTTGAAATCAATTTTG	803
Q	y	397	gcactcgagatgcttgaaggcgctctcttcggccctccatcgagagagggaatgatactta	456
D	b	804	GTCGTGAGATGATTTGGATTGGATTGGCTACCGAGTTTCATTAACCGTGGCATGTACTTGA	863
Q	y	457	tggatcgtggtgtgttaaccataatgagatgatgatgagcgggtgacgtccatgattaca	516
D	b	864	TGGTCGATATTGTTGTCATTCACCTTTGGCTTCTAGGAACCAATCCGACGTCGATTAAT	923
Q	y	517	gtgtgttaaacccgttcagtlcccaagaactctccacccgtctgttcttcaataaact	576
D	b	924	CTGAATATTTCATATTCGTGTCAGAGATTATTTCATTCAATTTTGGTGAATTCAGGATT	983
Q	y	637	atctcgataccaccaagggtgtgtgcaagaatgaaatgtacagcgggtggatcatctg	696
D	b	1044	ACGTTATATACACTTGGACACCGGTGAATAATGAAATATACGTCCCTGGGTATTAACACTTA	1103
Q	y	697	latoagaactatccatctgacgcgcctccgatatgacaaagataaacaacgtccagaagact	756
D	b	1104	TAGCTAATTACTGTATTGATGGTGTGATGAAATTGACACGTGTCAAGCATGTGAGATGGAATT	1163
Q	y	757	ctgtgcccgggtgtcaacaagaacgcgcgtgtactgtatcgcgagggtgtctcgaggtg	816
D	b	1164	TTTGGGGCCATTTCAGAGGCGACGAGGATTTAACGTGTGGTCAATATTCATGATG	1223
Q	y	817	atccgcgcctacactgtccctaccagaagctcatatgacgcgcgtactgaactatccatt	876
D	b	1224	ATTCATTCCTATATATGTCCGTATCAAGAAATCTGACGCGTGTGTAATTAATCCCGTTT	1283
Q	y	877	actatccactcctcacaagccttcaagtaactacccgcgcgaatcgtgacagcactctacaca	936
D	b	1284	ATTATCTCGTTGTCCTGGCTTTACACCGTGTGGTGGTGCATATCTCGTTAGTCGATA	1343
Q	y	937	tgtatcaaacgcgtcaaatccgcactgttcagagactcaaacactctctggacattcgtcgaga	996
D	b	1344	TGATTTGATACGCTCAAGTGTGAAATGCATGCACATCACTCTCTGGGGTCCTTTACAGA	1403
Q	y	997	accacgcacaacccaggttcgctcttcaacaacaagaatagccctcgcgaagaatcgc	1056
D	b	1404	ATTCAGGATATTCGCGAATTCCTCTTAACCTTCTGATGAGTCTTTAATTAATAATGGA	1463
Q	y	1057	cagatcatcatcctcctcaagaagaatcccatcatctacgcgcgcgaagaagaacact	1116
D	b	1464	TTGCTTTCACTATATCTCTGAGAGCGATTCCTATATATATTACGGCCCAAGACAGAC	1523
Q	y	1117	accgcgcgcgaagaacacccgcgcgaaccgcgaaacacactggtcttcgggtctaccgcagc	1176
D	b	1524	TCAATGCTGGAATAATGATTCCTATTAATTCGAGAGCGCTTTGGCTTCACGATTTCCACAA	1583
Q	y	1177	accagcgagctgatacaagtaattgtcctccgcgaaagcaatcccggaactatgcatatgaa	1236
D	b	1584	CTTCACGTTTACGATATACATGCGTCTGCTGAATTCAGATTCGAATATCTATATACA	1643
Q	y	1237	aagatatacagattcgtgnaactacaagaacttggccatctacaagaagacaacaagatcg	1296
D	b	1644	TAGATGATCTTATCTTACATATACAAATTTGGTTATTATTCCGATTTCCACGACTATAG	1703
Q	y	1297	ccatgycgaaggcaacagatggtgtcgacatcggtgactatctgttccaacaagggtgtct	1356
D	b	1704	CAATGCGGAAGGATTTACGGGGAAACCAATATATTACTGTTCTATCAATCTTGGGTACA	1763
Q	y	1357	cgggtgtatctgtataccctctctctgtagtggtgtggtgtgttaaccacgcgcgcgaagatg	1416
D	b	1764	GTCGCAATCTGTATCACTTTTACGCTTTCCAAATACGGGATATATCTCATCTATGCGTTGAT	1823

OY	1417	cggaggtattgctgcgcacgcacctgacggttgatcgatggaataagtgccttcccta	1476
Db	1824	ATGAGATCTTGACAAACGCACACTGTCACTGTGGATTGTCCGGAAATTGGCAGTCCAA	1883
OY	1477	tggcagggtggacttacccttgatgatgtatccgacaggaagttggcaggtagcaagctc	1536
Db	1884	TGTCGGGTTGGCTTACCAGAGATGTTTTACCCCGAATCGCACTGGTGTGCTGGAAATCT	1943
OY	1537	gtagtacgtcgtga	1550
Db	1944	GCTCATGTATAGAGA	1957
RESULT	15		
LOCUS	SCSMA2	2143 bp	DNA linear
DEFINITION	S. occidentalis SWA2 gene for alpha-amylose.		
ACCESSION	X73497		
VERSION	X73497.1		
KEYWORDS	GI:396561		
SOURCE	alpha-amyloses; SWA2 gene. Debaromyces occidentalis. Debaromyces occidentalis Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Debaromyces.		
ORGANISM			
REFERENCE	1 (bases 1 to 2143)		
AUTHORS	Claros,M.G., Abarca,D., Fernandez-Lobato,M. and Jimenez,A.		
TITLE	Molecular structure of the SWA2 gene encoding an AMY-related alpha-amyase from Schwanniomycos occidentalis		
JOURNAL	Curr. Genet. 24 (1-2), 75-83 (1993)		
MEDLINE	93365041		
REFERENCE	2 (bases 1 to 2143)		
AUTHORS	Jimenez,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-AUG-1993) A. Jimenez, Centro de Biologia Molecular, Universidad Autonoma, Cantoblanco, 28049 Madrid, SPAIN		
FEATURES			
Source	Location/Qualifiers		
	1..2143		
	/organism="Debaromyces occidentalis"		
	/strain="ARCC 26077"		
	/db_xref="taxon:27300"		
	234..1757		
gene	/gene="SWA2"		
	234..1757		
	/gene="SWA2"		
	/EC_number="3.2.1.1"		
	/codon_start=1		
	/product="alpha-amyase"		
	/protein_id="CAA51912.1"		
	/db_xref="GI:396562"		
	/db_xref="SPTRMBL:Q08806"		
	/translation="MKFATILSTALALSLVASKPIPLSKRDASAAAMRPSPTI OLVNDRAFRPGRGSATSACNTDGDPRYCGGTGGCITDKIDJIOGMFTAIWISPYVEQIF DDTGTGTAHGYTHMKMDIYAINSNGTLADLNLSNLELRKMKMLAVDIIVHYAMNGA GSSVAVSNNPFNOOSYPFDHCLITTNDQNVBDCWGVNTVSLPDRTSDSVSS FNLVWAELVSNYSIDGLRIDSAKHVDESFSPSAAGVYLIGEVYDGPAYTCPOYQ YMGSQTNPILYPMIREFOGTSNSYDELNAMISSLESCKDITLLGNFIENHDQRLFL SYTSDNALIKNAIFAENMISDGIPITYYGOEGYGSSDPNRRLAMLGISTSGNYRY LISSVNORNAITYKKDSKYTTYWSGVLYASGHVIAIQRGADQRIYSVFNNLGSSGS TVTSTKTYSGEKEVDVLTCTSTANDSTLTVSISSGAPRIIAPASLIANSIGNF"		
BASE COUNT	654 a	375 c	379 g
ORIGIN			735 t
Query Match	23 6%;	Score 409;	DB 8; Length 2143;
Best Local Similarity	58.4%;	Pred. No. 9.8e-86;	
Matches 715; Conservative	0;	Mismatches 510;	Indels 0; Gaps 0.
OY	121	ctgcggactggcgatcgcatcatattatcttcctcacgagtcgatttgcagaagcg	180
Db	337	CTGCAGACTTGGCGCTTCTGAATCTATCTACAACACTGTGTTCCGATAGATTTCGCAAGC	396
OY	181	atgggttcgcagactgcgacttgttaactcgcgatacgaataactgtgtggaacaatg	240

Db 397 ACGGATCGACTTCAGCTACTTGTAAATCTGSAATAGATATACTGTGGGGACTTTCC 456
Qy 241 agggcaatcagcaacttggactatatacaggaatggcttcaagcacttgatba 300
Db 457 AAGGTATTATGACAAATTTGGATTACATCCAAAGTATGGGTTCACTGCTATTGGATT 516
Qy 301 ccccgcttaagccagctgccccagaccacccatatagtagatgctctccatggtact 360
Db 517 CTCAGTGTGTGAACAATTCCTGATGATACCTGTATGTTATGCTTACACGGGTATT 576
Qy 361 ggcagcagatatatactctctgaacaaactaagcactgcaatgacttgaagcgc 420
Db 577 GGATGAAGATATTTACGCTATTAATTTCAATTTTGTACTCCGATGACTGGAAGATC 636
Qy 421 tctctcgccctcatgagagggtatctatctatgctgatgtgtgcttaaccata 480
Db 637 TTTCAAATGATTTGCAATGAAGAAATATGAAGCTTATGGTATGTTACTTAACCAT 696
Qy 481 tgggctatgatgagcgggtagctcagtcagttacagtgtgttaaccgttcagttccc 540
Db 697 ATGCTTGATGATGGTGGCGGAGAGCTGTGCTTACTCCAACTACATCCATTCACCAAC 756
Qy 541 aagaactacttcaacccgttctgttcatcnaaaactataaagatcagactcagttgag 600
Db 757 AATCTACTTCCACGATTTATTGTTTAAATTAACAATTAACATGATCAACCAATGTTGAAG 816
Qy 601 attgctgctagagataacacgtctcctctcgtatctcgtatcctacacaaagatggt 660
Db 817 ATTGCTGGGAAGGCGATTAATCTGTTATACAGATCTTCGACTGAGGATTCAGATG 876
Qy 661 tcaagaatgaatgtagactggtgtggaatcattgtatcgaactactccattgaacgcc 720
Db 877 TTAGCTCTATTTCATCTGCTGGTGGTGGAGTATGTTCTTAATTACTCAATTGATGTT 936
Qy 721 tccgtatcgaacagtaaaacacgltcagaaggacttctgcccgggtacaaagccg 780
Db 937 TAAGGATTCACAGCTAGCAGTGAAGCATGTAATTCATTCACCATTCACAAAGTGCTG 996
Qy 781 caagcggtactgactcggcgagggtctcgaacgtgataccggcctacactgtccctacc 840
Db 997 CAGGTCTATATCTTCTTGAGAAATTTATGACGGTGATCCAGCTTACACTTCCCATACC 1056
Qy 841 aagaactatgagcggcgtactgaactataccattactatccactcctcaacgcctca 900
Db 1057 AAACCTATATGTCAGGGGTTACTTAATCTTGTATCTATCCAAATGTAAGATCTTTC 1116
Qy 901 agtcaacctccggcagcatgagcagactctacaacatgatacaacgcgtcaaatccgact 960
Db 1117 AAGGTACTTCTAACCTGTCGATGAATTAATGCTATGATTTCAAGTTTGAAGATGATT 1176
Qy 961 gtccagactcaaacactcctcggcaactcgtcgaagaacagcaaccacaggttcgctt 1020
Db 1177 GTAAGATATTTACTTATTTGGGTAATTTTCATGTAACCATGATCAACCAAGATTACAT 1236
Qy 1021 cttaacaacaagacatagccctcgcgaagaacgtcgacgaltcatcctcaacgacg 1080
Db 1237 CTTATCTTCTGATAGTGCTTAATCAAAAATGCAATTGCTTATTTAATGTCAGATG 1296
Qy 1081 gaatcccatctatcagccggccaagaacagcaactacgcccggaaagaccccgca 1140
Db 1297 GATATTCAAATTTATTTACTACGGGTCAAGAACAGGTTACAGTGTAGCTCCGATCCAAACA 1356
Qy 1141 accgcgaagcaaacctgctcctcgggtacccgacagcagcgtgtacaagttaatgt 1200
Db 1357 ACAGAGAAACATTTATGTTATCTGTTACAGCATAGTATGTTACTTACAACTTATCT 1416
Qy 1201 cctcgcggaacgaatccggaactatgcccattagaagaatatacagatcgtgacctaca 1260
Db 1417 CTTCAATTTAATTAACAACCAAGCATTTTATAGGATAGCAAAATACACTACTTATT 1476
Qy 1261 agaactgcccactctacaagaagacaaacgactcgcacatgccaaggccacagatgggt 1320

Db 1477 GGAGTATGTTATTAACGCTTCAGGTCAATGTTATGCTTCAAAAGAGTGACAGCACC 1536
Qy 1321 cgcagatcgttactatcttgcaca 1345
Db 1537 AAAGATTTGTTCTGTCTTTAACAA 1561

Search completed: June 13, 2002, 18:23:34
Job time: 16672 sec

Fri Jun 14 10:36:35 2002

us-09-710-339-1.rge

Page 25

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 13, 2002, 08:54:29 ; Search time 57.95 Seconds
(without alignments)
4.636 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160
Perfect score: 57
Sequence: 1 SSVDSYFKPF 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents.AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTOS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	478	1	US-08-720-899-7 Sequence 7, Appli
2	57	100.0	478	1	US-08-459-610-7 Sequence 7, Appli
3	57	100.0	478	1	US-08-343-804-7 Sequence 7, Appli
4	57	100.0	478	2	US-08-339-715A-2 Sequence 2, Appli
5	57	100.0	478	2	US-08-600-908A-10 Sequence 10, Appli
6	57	100.0	478	3	US-08-683-838A-10 Sequence 10, Appli
7	57	100.0	478	4	US-09-182-859-7 Sequence 7, Appli
8	37	64.9	468	1	US-08-204-656B-2 Sequence 2, Appli
9	37	64.9	468	1	US-08-204-656B-4 Sequence 4, Appli
10	37	64.9	468	1	US-08-204-656B-6 Sequence 6, Appli
11	37	64.9	468	1	US-08-204-656B-8 Sequence 8, Appli
12	37	64.9	468	1	US-08-470-702-6 Sequence 6, Appli
13	37	64.9	468	1	US-08-470-702-7 Sequence 7, Appli
14	37	64.9	468	1	US-08-470-702-8 Sequence 8, Appli
15	37	64.9	468	1	US-08-467-831-6 Sequence 6, Appli
16	37	64.9	468	1	US-08-467-831-7 Sequence 7, Appli
17	37	64.9	468	1	US-08-467-831-8 Sequence 8, Appli
18	37	64.9	468	1	US-08-467-831-9 Sequence 9, Appli
19	37	64.9	468	1	US-09-105-390-48 Sequence 48, Appli
20	35	61.4	377	4	US-09-105-390-64 Sequence 64, Appli
21	35	61.4	497	4	US-09-058-947A-4 Sequence 4, Appli
22	35	61.4	500	4	US-08-868-373-12 Sequence 12, Appli
23	33	57.9	1507	6	US-08-296-791-3 Sequence 3, Appli
24	33	57.9	1541	4	US-08-296-791-3 Sequence 3, Appli
25	33	57.9	1541	5	US-08-296-791-3 Sequence 3, Appli
26	33	57.9	1541	5	US-08-296-791-3 Sequence 3, Appli
27	33	57.9	1545	4	US-08-296-791-4 Sequence 4, Appli

28	33	57.9	1545	5	PCT-US95-10661A-4 Sequence 4, Appli
29	33	57.9	1702	4	US-08-296-791-5 Sequence 5, Appli
30	33	57.9	1702	5	PCT-US95-10661A-5 Sequence 5, Appli
31	33	57.9	1848	4	US-08-296-791-6 Sequence 6, Appli
32	33	57.9	1848	5	PCT-US95-10661A-6 Sequence 6, Appli
33	32	56.1	172	1	US-08-487-001A-51 Sequence 51, Appli
34	32	56.1	172	2	US-08-630-822A-51 Sequence 51, Appli
35	32	56.1	172	2	US-09-005-069-51 Sequence 51, Appli
36	32	56.1	182	2	US-08-874-832-14 Sequence 14, Appli
37	32	56.1	182	2	US-09-097-233-14 Sequence 14, Appli
38	32	56.1	213	2	US-08-874-832-5 Sequence 5, Appli
39	32	56.1	213	3	US-09-097-233-5 Sequence 5, Appli
40	32	56.1	425	4	US-09-247-155-99 Sequence 99, Appli
41	32	56.1	510	1	US-08-255-670A-2 Sequence 2, Appli
42	32	56.1	675	1	US-08-317-822A-9 Sequence 9, Appli
43	32	56.1	675	1	US-08-439-818A-9 Sequence 9, Appli
44	32	56.1	675	2	US-08-751-965-9 Sequence 9, Appli
45	32	56.1	675	2	US-08-738-975-9 Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-720-899-7
Sequence 7, Application US/08720899
Patent No. 5753460
GENERAL INFORMATION:
APPLICANT: Bisgaard-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
APPLICANT: Svendsen, Allan
APPLICANT: Thellersen, Marianne
APPLICANT: Van der Zee, Pia
TITLE OF INVENTION: AMYLASE VARIANTS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,899
FILING DATE: 10-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/343,804
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054,214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-720-899-7
Query Match 100.0%; Score 57; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
1111111111
Db 130 SSVDSVFKPF 140

RESULT 2

US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-459-610-7

Query Match 100.0%; Score 57; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
1111111111
Db 130 SSVDSVFKPF 140

RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58308370 No. 5830837disk of No. 5830837th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-343-804-7

Query Match 100.0%; Score 57; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
1111111111
Db 130 SSVDSVFKPF 140

RESULT 4
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michio
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; STREET: G. Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Melser, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 57; DB 2: Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
DB 130 SSVDSYVFKPF 140

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-600-908A-10

Query Match 100.0%; Score 57; DB 2: Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
DB 130 SSVDSYVFKPF 140

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-683-838A-10

Query Match 100.0%; Score 57; DB 3: Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
DB 130 SSVDSYVFKPF 140

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgard-Frantzen, Henrik

TITLE OF INVENTION: Alpha-Amylase Mutants
FILE REFERENCE: 4796.204-US
CURRENT APPLICATION NUMBER: US/09/182,859
CURRENT FILING DATE: 1998-10-29
EARLIER APPLICATION NUMBER: 0515/96
EARLIER FILING DATE: 1996-04-30
EARLIER APPLICATION NUMBER: 0712/96
EARLIER FILING DATE: 1996-06-28
EARLIER APPLICATION NUMBER: 0775/96
EARLIER FILING DATE: 1996-07-11
EARLIER APPLICATION NUMBER: 1263/96
EARLIER FILING DATE: 1996-11-08
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 7
LENGTH: 478
TYPE: PRT
ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 57; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDYSEVFKPF 11
Db 130 SVDYSEVFKPF 140

RESULT 8

US-08-204-656B-2
Sequence 2, Application US/08204656B

Patent No. 5538882

GENERAL INFORMATION:

APPLICANT: Matsui, Ikuro

APPLICANT: Ishikawa, Kazuhiko

APPLICANT: Miyairi, Sachio

APPLICANT: Honda, Koichi

TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,

TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP

STREET: 8110 Gatehouse Road, Suite 500 East

CITY: Falls Church

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22042

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/204,656B

FILING DATE: 02-MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Weiner, Marc S.

REGISTRATION NUMBER: 32,181

REFERENCE/DOCKET NUMBER: 234-252P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

TELEX: 248345

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 468 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-204-656B-2

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SVDYSEVFKPF 11
Db 132 SVDYSEVFKPF 141

RESULT 9

US-08-204-656B-4
Sequence 4, Application US/08204656B

Patent No. 5538882

GENERAL INFORMATION:

APPLICANT: Matsui, Ikuro

APPLICANT: Ishikawa, Kazuhiko

APPLICANT: Miyairi, Sachio

APPLICANT: Honda, Koichi

TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,

TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP

STREET: 8110 Gatehouse Road, Suite 500 East

CITY: Falls Church

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22042

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/204,656B

FILING DATE: 02-MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Weiner, Marc S.

REGISTRATION NUMBER: 32,181

REFERENCE/DOCKET NUMBER: 234-252P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

TELEX: 248345

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 468 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-204-656B-4

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SVDYSEVFKPF 11
Db 132 SVDYSEVFKPF 141

RESULT 10

US-08-204-656B-6
Sequence 6, Application US/08204656B

Patent No. 5538882

GENERAL INFORMATION:

APPLICANT: Matsui, Ikuro

APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Miyairi, Sachio
APPLICANT: Honda, Koichi
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
TITLE OF INVENTION: Oligosaccharide Using The Enzyme
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-204-656B-6

Query Match 64.9%; Score 37; DB 1; Length 468;
Best local similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SYDYSVFKPF 11
1:111:11
DB 132 SIDYSEYTF 141

RESULT 11
US-08-204-656B-8
Sequence 8, Application US/08204656B
Patent No. 5538882
GENERAL INFORMATION:
APPLICANT: Matsui, Ikuro
APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Miyairi, Sachio
APPLICANT: Honda, Koichi
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
TITLE OF INVENTION: Oligosaccharide Using The Enzyme
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-204-656B-8

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-204-656B-8

Query Match 64.9%; Score 37; DB 1; Length 468;
Best local similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SYDYSVFKPF 11
1:111:11
DB 132 SIDYSEYTF 141

RESULT 12
US-08-470-702-6
Sequence 6, Application US/08470702
Patent No. 5651149
GENERAL INFORMATION:
APPLICANT: MATSUI, IKURO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-6

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 SVDYSVEKPF 11
1:111:11
Db 132 SIDYSEYTFP 141

RESULT 13

US-08-470-702-7
; Sequence 7, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-470-702-7

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 SVDYSVEKPF 11
1:111:11
Db 132 SIDYSEYTFP 141

RESULT 14

US-08-470-702-8
; Sequence 8, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-470-702-8

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 SVDYSVEKPF 11
1:111:11
Db 132 SIDYSEYTFP 141

RESULT 15

US-08-470-702-9
; Sequence 9, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:

APPLICANT: MATSUI, IKUO
 APPLICANT: ISHIKAWA, KAZUHIKO
 APPLICANT: MIYAIRI, SACHIO
 APPLICANT: HONDA, KOICHI
 TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
 TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
 TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 STREET: 810 Gatehouse Road, Suite 500 East
 City: Falls Church
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22042
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentia Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/470,702
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/204,656
 FILING DATE: 02-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: WEINER, MARC S.
 REGISTRATION NUMBER: 32,181
 REFERENCE/DOCKET NUMBER: 234-252P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 TELEX: 248345
 INFORMATION FOR SEQ. ID NO. 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 468 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-470-702-9

Query Match 64.9%; Score 37; DB 1; Length 468;
 Best local Similarity 60.0%; Pred. No. 27;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 SVDSVFKPF 11
 1:111:11
 DB 132 SIDSEYTPF 141

Search completed: June 13, 2002, 08:54:30
 Job time: 315 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2002, 09:11:07 ; Search time 125.81 seconds
(without alignments)
15.126 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160

Perfect score: 57

Sequence: 1 SSVDYSVFKPF 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_IVIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	57	100.0	499	3	Q96TH4	Q96TH4 aspergillus
2	45	78.9	490	3	Q9UV07	Q9UV07 emericella
3	44	77.2	640	3	O13296	O13296 aspergillus
4	39	68.4	923	6	097745	097745 sus scrofa
5	39	68.4	945	6	077589	077589 equus caball
6	39	68.4	964	6	097744	097744 sus scrofa
7	39	68.4	964	6	Q9T001	Q9T001 sus scrofa
8	39	68.4	978	6	Q9XS93	Q9XS93 canis famli
9	39	68.4	978	6	Q9XS93	Q9XS93 canis famli
10	39	68.4	4116	5	09XPD2	09XPD2 drosophila
11	39	68.4	5322	5	09XPD2	09XPD2 drosophila
12	39	68.4	5322	5	09XPD2	09XPD2 drosophila
13	38	66.7	350	10	Q9FGE2	Q9FGE2 arabidopsis
14	38	66.7	425	16	Q9KRR9	Q9KRR9 vibrio chol
15	38	66.7	430	3	O13699	O13699 schizosacch
16	38	66.7	623	3	Q9UV09	Q9UV09 emericella

17	38	66.7	1869	3	Q93809	Q93809 magnaporthe
18	38	66.7	2160	5	Q9U5P3	Q9U5P3 halictis tu
19	38	66.7	2500	5	Q9NDV3	Q9NDV3 halictis tu
20	37	64.9	3419	5	Q9GP18	Q9GP18 halictis tu
21	37	64.9	245	8	002665	002665 podopora a
22	37	64.9	790	4	Q96MR9	Q96MR9 homo sapien
23	37	64.9	4820	5	Q9VAV5	Q9VAV5 drosophila
24	36	63.2	112	10	Q9LZ08	Q9LZ08 arabidopsis
25	36	63.2	173	5	Q9XNM1	Q9XNM1 caenorhabdi
26	36	63.2	257	16	Q98IS0	Q98IS0 rhizobium 1
27	36	63.2	300	16	Q9CGG9	Q9CGG9 lactococcus
28	36	63.2	353	5	Q9XVC7	Q9XVC7 caenorhabdi
29	36	63.2	409	10	Q9SCU1	Q9SCU1 arabidopsis
30	36	63.2	491	3	013996	013996 schizosacch
31	36	63.2	499	3	Q9UT42	Q9UT42 schizosacch
32	36	63.2	541	2	Q9ZNC7	Q9ZNC7 clostridium
33	36	63.2	579	11	Q9WV10	Q9WV10 mus muscui
34	36	63.2	1301	4	Q96GR7	Q96GR7 homo sapien
35	36	63.2	1329	4	Q9HCR3	Q9HCR3 homo sapien
36	36	63.2	2692	5	Q21547	Q21547 caenorhabdi
37	35.5	62.3	690	16	Q98I18	Q98I18 rhizobium 1
38	35	61.4	143	8	Q9XNE1	Q9XNE1 eriococcus
39	35	61.4	363	1	Q9UWQ2	Q9UWQ2 pyrococcus
40	35	61.4	377	10	Q9ZNY6	Q9ZNY6 oryza sativ
41	35	61.4	428	5	Q24024	Q24024 drosophila
42	35	61.4	428	5	Q9VFR3	Q9VFR3 drosophila
43	35	61.4	483	4	Q9H798	Q9H798 homo sapien
44	35	61.4	492	10	Q9C6L5	Q9C6L5 arabidopsis
45	35	61.4	497	10	Q9XPF3	Q9XPF3 arabidopsis

ALIGNMENTS

RESULT	ID	Q96TH4	PRELIMINARY:	PRT:	499 AA.
AC	Q96TH4	01-DEC-2001 (TRMBUREL.19, Created)			
DT	Q96TH4	01-DEC-2001 (TRMBUREL.19, Last sequence update)			
DT	01-DEC-2001	(TRMBUREL.19, Last annotation update)			
DE	TAKA-AMYLASE A (EC 3.2.1.1)	(ALPHA-AMYLASE).			
GN	AMYA OR AMY1.				
OS	Aspergillus oryzae, and				
OC	Aspergillus flavus.				
CC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_TaxID=5062, 5059;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=A.oryzae; STRAIN=RI840;				
RX	MEDLINE=20289310; PubMed=10830498;				
RA	Goni K., Akono T., Mineloki T., Ozeki K., Kumagai C., Okazaki N.,				
RA	Imura Y.;				
RT	"Molecular cloning and characterization of a transcriptional activator				
RT	gene, amyR, involved in the amylolytic gene expression in Aspergillus				
RT	oryzae.";				
RL	Biosci. Biotechnol. Biochem. 64:816-827(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=A.flavus; STRAIN=86-10D;				
RA	Fakhoury A.M., Woloshuk C.P.;				
RT	"Amy1, the alpha-amylase gene of Aspergillus flavus: Involvement in				
RT	atlatoxin biosynthesis in maize kernels.";				
RL	Phytopathology 89:908-914(1999).				
DR	EMBL; AB021876; BA95703.1; -.				
DR	EMBL; AF139925; AA14264.1; -.				
SQ	SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;				

Query Match 100.0%; Score 57; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0091;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
 |||||
 Db 151 SSVDSYVFKPF 161

RESULT 2
 ID Q9UV07 PRELIMINARY; PRT; 490 AA.
 AC Q9UV07;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ALPHA-AMYLASE AMYA.
 GN AMYA.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OC NCBITaxID=5072;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Boase N.A., Murphy R.M., Kelly J.M.;
 RT "An amylase cluster in Aspergillus nidulans."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF208225; AAF17103.1; -.
 DR HSSP; P10529; 7TAA.
 DR InterPro; IPR000461; Alpha-amylase.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR Pfam; PF02806; alpha-amylase; C; 1.
 DR PRINTS; PR00110; ALPHAMYLASE.
 SO SEQUENCE 490 AA; 54249 MW; A891C4ACEAE5305 CRC64;

Query Match 78.9%; Score 45; DB 3; Length 490;
 Best Local Similarity 72.7%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
 :|||||
 Db 143 ASVDYSAENPF 153

RESULT 3
 ID 013296 PRELIMINARY; PRT; 640 AA.
 AC 013296;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ACID-STABLE ALPHA-AMYLASE.
 OS Aspergillus kawachii.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OC NCBITaxID=40384;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Karako A.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DT [2]
 RT "Molecular-cloning and determination of the nucleotide-sequence of a gene encoding an acid-stable alpha-amylase from Aspergillus-kawachii."
 RL J. Ferment. Bioeng. 81:292-298(1996).
 DR EMBL; AB008370; BAA22993.1; -.
 DR HSSP; P56271; 2AAA.
 DR InterPro; IPR000461; Alpha-amylase.
 DR InterPro; IPR002044; CBD_4.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR Pfam; PF00686; CBD_4; 1.
 DR ProDom; PD001568; CBD_4; 1.
 SO SEQUENCE 640 AA; 69507 MW; 676BF9D0236720DD CRC64;

Query Match 77.2%; Score 44; DB 3; Length 640;
 Best Local Similarity 72.7%; Pred. No. 3.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
 :|||||
 Db 151 NDVDSYVFDPF 161

RESULT 4
 ID 097745 PRELIMINARY; PRT; 923 AA.
 AC 097745;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
 GN KIT2*0201.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC NCBITaxID=9823;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
 RA Marklund S.;
 RL Theses (1997), Department of Animal Breeding and Genetics,
 RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
 RA Marklund S.;
 RX MEDLINE=98391767; PubMed=9724328;
 RA Marklund S., Kijas U., Rodriguez-Martinez H., Roennstrand L., Funa K.,
 RA Moller M., Editors-Lilja I., Andersson L.;
 RT "Molecular evolution of the dominant white phenotype in the domestic pig."
 RL Genome Res. 8:826-833(1998).
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.

DR EMBL; AJ223231; CAI1199.1; -.
 DR HSSP; P11362; 1FGK.
 DR InterPro; IPR00719; Euk_pkinase.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003600; Ig_like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR001824; Receptor_tyr_kin_III.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00047; Ig_2.
 DR Pfam; PF00069; pkinase; 3.
 DR PRINTS; PR00109; TYRKINASE.
 DR SMART; SM00409; IG_1.
 DR SMART; SM00410; IG_like; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ARP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_ILI; 1.
 KW APP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
 KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
 FT NON_TER 923
 SO SEQUENCE 923 AA; 103780 MW; FDCD566F5D86B9F CRC64;

Query Match 68.4%; Score 39; DB 6; Length 923;
 Best Local Similarity 77.8%; Pred. No. 53;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDSYVFK 9
 :|||||
 Db 476 SSIDYSAFK 484

```

RESULT 5
ID 077589 PRELIMINARY; PRT: 945 AA.
AC 077589;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DE 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE TYROSINE KINASE RECEPTOR HOMOLOG (FRAGMENT).
GN KIT.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Lear T.L., Adams M.H., McDowell K.J., Ponce de Leon F.A.,
RA Sullivan N.D., Coogee L., Ferguson E., Chambers T.M., Bailey E.;
RT "Chromosomal location of genes for ESR, ETS2, G0T2, KIT, MXI and PGR
RT in the horse, Equus caballus."
RT Proc. Xth Amer. Coll. Dom. Anim. Cytogenet. Gene Mapp. 0:0-0(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;
RT "An Equine sequence homologous to the c-KIT proto-oncogene mapped to
RT chromosome 3q13-21."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AF055037; AAC23611.1; -.
DR HSSP: P11362; IFGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; Receptor_tyr_kin_III.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; Pkinase; 2.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00409; IG; 2.
DR SMART: SM00219; TYKIC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
FT NON_TER 945
SQ SEQUENCE 945 AA; 106143 MW; 7DB424A89DE5C340 CRC64;

```

```

Query Match 68.4%; Score 39; DB 6; Length 945;
Best Local Similarity 77.8%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 SSVDSYVK 9
DB 476 SSIDYSAFK 484

```

```

RESULT 6
O97744 PRELIMINARY; PRT: 964 AA.
AC O97744;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
GN KIT1*0101.

```

```

OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RX MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funä K.,
RA Moller M., Editors-Lilla I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig."
RL Genome Res. 8:826-833(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AJ223228; CA01196.1; -.
DR HSSP: P11362; IFGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; Receptor_tyr_kin_III.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; Pkinase; 2.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00409; IG; 1.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00219; TYKIC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
FT NON_TER 964
SQ SEQUENCE 964 AA; 108287 MW; 90E72EDFAB1358B7 CRC64;

```

```

Query Match 68.4%; Score 39; DB 6; Length 964;
Best Local Similarity 77.8%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 SSVDSYVK 9
DB 476 SSIDYSAFK 484

```

```

RESULT 7
O97001 PRELIMINARY; PRT: 964 AA.
AC O97001;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
GN KIT1*0201.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,

```

RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
RA MEDLINE=98391767; PubMed=9724328;
RX Marklund S., Kljas J., Rodriguez-Martinez H., Roenstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
pig."
RL Genome Res. 8:826-833(1998).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AJ223229; CAA11197.1; -.
DR HSSP: P11362; 1FGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; Receptor_tyr_kin_III.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; pkinase; 2.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00409; IG; 1.
DR SMART: SM00410; IG_like; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
FT NON_TER 964 964
SQ SEQUENCE 964 AA; 108315 MW; 996C3C46201356A8 CRC64;

Query Match 68.4%; Score 39; DB 6; Length 964;
Best Local Similarity 77.8%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDSVFK 9
||:|||||
DB 476 SSIDYSAFK 484

RESULT 8
O9TQO0 PRELIMINARY; PRT; 964 AA.
AC O9TQO0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
GN K11*0202.
OS Sus scrofa (Pig).
OC Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_Taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
RX MEDLINE=98391767; PubMed=9724328;
RX Marklund S., Kljas J., Rodriguez-Martinez H., Roenstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
pig.";

RL Genome Res. 8:826-833(1998).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AJ223230; CAA11198.1; -.
DR HSSP: P11362; 1FGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; Receptor_tyr_kin_III.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; pkinase; 2.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00409; IG; 1.
DR SMART: SM00410; IG_like; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
FT NON_TER 964 964
SQ SEQUENCE 964 AA; 108287 MW; 1B21292A962B9191 CRC64;

Query Match 68.4%; Score 39; DB 6; Length 964;
Best Local Similarity 77.8%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDSVFK 9
||:|||||
DB 476 SSIDYSAFK 484

RESULT 9
O9XS93 PRELIMINARY; PRT; 978 AA.
AC O9XS93;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE KIT.
GN C-KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBELLUM;
RA London C.A., Galli S.J., Yuuki T., Hu Z.-Q., Helfand S.C.,
RA Geisler E.N.;
RT "Spontaneous canine mast cell tumors express tandem duplications in
the proto-oncogene c-kit.";
RL Exp. Hematol. 0:0-0(1999).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AF099030; AAD28369.1; -.
DR HSSP: P11362; 1FGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; Receptor_tyr_kin_III.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; pkinase; 2.
DR SMART: SM00409; IG; 2.
DR SMART: SM00410; IG_like; 1.

DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
 SQ Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
 SEQUENCE 978 AA; 109651 MW; 5AC31E5AC4E9910F CRC64;

Query Match 68.4%; Score 39; DB 6; Length 978;
 Best Local Similarity 77.8%; Pred. No. 56;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDYSVFK 9
 ||:|||||
 DB 478 SSIDYSAFK 486

RESULT 10
 O96958 PRELIMINARY; PRT; 4116 AA.

AC O96958: 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE CALO PROTEIN (FRAGMENT).
 GN POE OR CALO OR CG14472.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RETINAL;
 RX MEDLINE=99030403; PubMed=9813038;
 RA Xu X.Z.S., Wes P.D., Chen H., Li H.S., Yu M., Morgan S., Liu Y.,
 RA Montell C.;
 RT "Retinal targets for calmodulin include proteins implicated in
 RT synaptic transmission";
 RL J. Biol. Chem. 273:31297-31297(1998).
 DR EMBL: Y17920; CAA76940.1; -;
 DR FlyBase: FBgn0011230; Poe.
 DR InterPro: IPR002106; AA.Trna.Ligase_II.
 DR InterPro: IPR003126; Znf_UBR1.
 DR Pfam: PF02207; zf-UBR1; 1.
 DR SMART: SM00396; Znf_UBR1; 1.
 DR PROSITE: PS00339; AA.Trna.Ligase_II_2; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 4116 AA; 457563 MW; D3DFB691E7EF3F36 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 4116;
 Best Local Similarity 75.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 DYSVFKPF 11
 ||:|||||
 DB 4013 DYTIVKPF 4020

RESULT 11
 O9XYD2 PRELIMINARY; PRT; 5322 AA.

AC O9XYD2: 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PUSHOVER.
 GN POE OR PUSH OR CG14472.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Richards S., Hekmat-Scafe D.S., Hurd D.D., Caprette D.R., Saxton W.M.,
 RA Carlson J.R., Stern M.J.;
 RT "Hyperplastic glial cells surrounding abdominal nerves in pushover and
 RT irradiated pushover mutants of Drosophila melanogaster";
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF096897; AAD20450.1; -;
 DR FlyBase: FBgn0011230; Poe.
 DR InterPro: IPR002106; AA.Trna.Ligase_II.
 DR InterPro: IPR003126; Znf_UBR1.
 DR Pfam: PF02207; zf-UBR1; 1.
 DR SMART: SM00396; Znf_UBR1; 1.
 DR PROSITE: PS00339; AA.Trna.Ligase_II_2; UNKNOWN_1.
 FT VARIANT 728 728 Q -> *
 FT VARIANT 883 883 Q -> *
 SQ SEQUENCE 5322 AA; 590679 MW; ABP4C4BC7A9CEBFE CRC64;

Query Match 68.4%; Score 39; DB 5; Length 5322;
 Best Local Similarity 75.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 DYSVFKPF 11
 ||:|||||
 DB 5219 DYTIVKPF 5226

RESULT 12
 O9VLT5 PRELIMINARY; PRT; 5322 AA.

AC O9VLT5: 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE CG14472 PROTEIN.
 GN POE OR CG14472.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,
 RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Aghayani A., An H.-J., Andrews-Plamkuch C., Baldwin D.,
 RA Balieu R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhargava D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Ditzel S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Styrekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Glibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "the genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003619; AAF52598.1; -.
 DR FLYBase: FBgn0011230; poe.
 DR InterPro: IPR002106; AA:trna_ligase_II.
 DR InterPro: IPR003126; znf.UBR1.
 DR Pfam: PF02207; zf.UBR1.1.
 DR SMART: SM00396; znf.UBR1.1.
 DR PROSITE: PS00339; AA:trna_ligase_II.2; UNKNOWN.1.
 DR SEQUENCE 5322 AA; 590696 MW; 3890824E00F2DB50 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 5322;
 Best Local Similarity 75.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 4 DYSVFKPF 11
 DB 5219 DYTYYKPF 5226

RESULT 13
 Q9FGE2 PRELIMINARY; PRT; 350 AA.
 AC Q9FGE2;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE GBAAD32930.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
 RL Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB025641; BAB08931.1; -.
 DR SEQUENCE 350 AA; 39471 MW; C33011C2DB86C2DB CRC64;

Query Match 66.7%; Score 38; DB 10; Length 350;
 Best Local Similarity 54.5%; Pred. No. 29;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SSVDSVFKPF 11
 DB 329 SSMYKSYKPF 339

RESULT 14
 Q9KRR9 PRELIMINARY; PRT; 425 AA.
 AC Q9KRR9;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE HYPOTHETICAL PROTEIN VC1567.
 GN VC1567.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OC NCBI_TaxID=666;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodsan R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004234; AAF94721.1; -.
 DR TIGR: VC1567;
 DR InterPro: IPR003838; DUF214.
 DR Pfam: PF02687; DUF214.1.
 DR Complete proteome.
 DR SEQUENCE 425 AA; 46328 MW; 982D2C2DBAFA9EE7 CRC64;

Query Match 66.7%; Score 38; DB 16; Length 425;
 Best Local Similarity 63.6%; Pred. No. 36;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SSVDSVFKPF 11
 DB 209 SPIDYSGFTPF 219

RESULT 15
 ID O13699 PRELIMINARY; PRT; 430 AA.
 AC O13699;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
 DE HYPOTHETICAL 50.4 KDA PROTEIN C1167.06C IN CHROMOSOME I.
 GN SPAC1167.06C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxID=4896;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA McLean J., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
 RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 CC -I- SIMILARITY: SOME, TO THE C-TERMINAL OF S.POMBE SPAC469.07.
 DR EMBL: Z99161; CAB16211.1; -.
 DR KW Hypothetical protein.
 DR SEQUENCE 430 AA; 50433 MW; 6FE023D9545B286A CRC64;

Query Match 66.7%; Score 38; DB 3; Length 430;
 Best Local Similarity 70.0%; Pred. No. 36;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 SVDYSVFKPF 11
 DB 21 SVDYSQFKAF 30

Search completed: June 13, 2002, 09:11:10
 Job time: 1055 sec

Fri Jun 14 10:36:43 2002

us-09-710-339-2_copy_150_160.rsp

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2002, 09:11:49 ; Search time 38.1 Seconds

(without alignments)
11.179 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160

Perfect score: 57

Sequence: 1 SSVDYSVFKPF 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	498	1	AMYA_ASPAW
2	57	100.0	499	1	AMYA_ASPOR
3	57	100.0	499	1	AMYA_ASPAW
4	57	100.0	499	1	AMYA_ASPSH
5	45	78.9	512	1	AMYA_DEBOC
6	44	77.2	484	1	AMYA_ASPNG
7	39	68.4	975	1	KIT_CANPA
8	39	68.4	978	1	KIT_CANPA
9	38	66.7	513	1	AMT2_SCHPO
10	37	64.9	494	1	AMT1_SACFI
11	36	63.2	521	1	DRTS_MAIZE
12	36	63.2	1367	1	DPO3_THEMA
13	35	61.4	215	1	IF4E_APLCA
14	35	61.4	409	1	ARCA_BORAF
15	35	61.4	410	1	ARCA_BORAF
16	35	61.4	410	1	ARCA_BORAF
17	34	59.6	1390	1	HCTB_HELPD
18	34	59.6	1391	1	N155_RAT
19	33	57.9	200	1	N155_HUMAN
20	33	57.9	260	1	YQAB_HAEIN
21	33	57.9	277	1	YQ56_METUA
22	33	57.9	334	1	SP0A_CLOIN
23	33	57.9	352	1	CUB_CLOIN
24	33	57.9	352	1	FBRL_CABEL
25	33	57.9	356	1	GATA_TREPA
26	33	57.9	525	1	VE59_LAMBD
27	33	57.9	530	1	DRIS_SOYBN
28	33	57.9	539	1	YQ33_METUA
29	33	57.9	561	1	ILVD_METUA
30	33	57.9	607	1	YJRO_YEAST
31	33	57.9	685	1	YQ04_YEAST
32	33	57.9	707	1	ETP2_MCVI
33	33	57.9	710	1	ETP2_VACCA
33	33	57.9	710	1	ETP2_VACCOC

34	33	57.9	710	1	ETP2_VACCT	O9J1F93 vaccinia vi
35	33	57.9	710	1	ETP2_VACCV	P20636 vaccinia vi
36	33	57.9	710	1	ETP2_VARV	P33806 variola vir
37	33	57.9	711	1	ETP2_MXVL	O968K4 myxoma vir
38	33	57.9	711	1	ETP2_SPVKA	O968Y2 shope fibro
39	33	57.9	713	1	ETP2_YRBAW	O968B4 yaba monkey
40	33	57.9	775	1	POP1_SCHPO	P87060 schizosacch
41	33	57.9	1195	1	HEUS_METUA	O58524 methanococ
42	33	57.9	1458	1	CUT4_SCHPO	O9urv2 schizosacch
43	33	57.9	1532	1	IGA_NEIGO	P093790 neisseria g
44	33	57.9	1541	1	IGA1_HAEIN	P42782 haemophilus
45	33	57.9	1545	1	IGA3_HAEIN	P45385 haemophilus

ALIGNMENTS

RESULT 1	AMYA_ASPAW	STANDARD;	PRT;	498 AA.
AC	002905;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan			
DE	glucanohydrolase A).			
GN	AMYA.			
OS	Aspergillus awamori.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=105351;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRATN-DVK143F;			
RX	MEDLINE=90254827; PubMed=2340591;			
RA	Korman D.R., Bayliss F.T., Barnett C.C., Cartmona C.L.,			
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,			
RA	Berka R.M.;			
RT	"Cloning, characterization, and expression of two alpha-amylase genes			
RT	from Aspergillus niger var. awamori."			
RU	Curr. Genet. 17:203-212(1990).			
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic			
CC	linkages in oligosaccharides and polysaccharides.			
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO			
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-slb.ch/announce/			
CC	or send an email to license@isb-slb.ch).			
CC	-----			
DR	EMBL; X52755; CA36966.1; -.			
DR	HSSP; P10529; TTA.			
DR	InterPro; IPR000461; Alpha_amylase.			
DR	Pfam; PF00128; alpha-amylase; 1.			
KW	Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;			
KW	Glycoprotein; signal; Multigene family.			
FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	498	ALPHA-AMYLASE A.
FT	ACT_SITE	227	227	BY SIMILARITY.
FT	ACT_SITE	231	231	BY SIMILARITY.
FT	ACT_SITE	251	251	BY SIMILARITY.
FT	ACT_SITE	318	318	BY SIMILARITY.
FT	DISULFID	51	59	BY SIMILARITY.
FT	DISULFID	171	185	BY SIMILARITY.
FT	DISULFID	261	304	BY SIMILARITY.
FT	DISULFID	461	496	BY SIMILARITY.
FT	CARBOHYD	218	218	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SEQUENCE	498 AA;	54880 MW;	7658511BC01A8A01 CRC64;

Query Match 100.0%; Score 57; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.0011;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVDSYVFKEP 11
 |||||
 DB 151 SSVDSYVFKEP 161

RESULT 2

AMYL ASPOR STANDARD; PRT; 499 AA.
 ID AMYL ASPOR STANDARD; PRT; 499 AA.
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 DE (1,4-alpha-D-glucan glucanohydrolase).
 GN AMY1 AND AMY2 AND AMY3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 63303;
 RX MEDLINE=89237897; PubMed=2785629;
 RA Mirel S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT intron-exon organization";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89378767; PubMed=2789162;
 RA Genes M.J., Dove M.J., Seliy V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90128276; PubMed=2612911;
 RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
 RA Ueki S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A";
 RL Proc. Jpn. Acad. B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RX MEDLINE=74001521; PubMed=4733850;
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RT A with trypsin and chymotrypsin";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RL Proc. Jpn. Acad. B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; PubMed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84212370; PubMed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; PubMed=9283074;
 RA Brozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution";
 RL Biochemistry 36:10837-10845(1997).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04548; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: JY0466; JY0466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR Glycosultdb: P10529; -;
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR HydroLase: glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT STGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT VARIANT 93 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTID=CAR_000123.
 Q -> R (IN AMY3).
 F -> L (IN AMY3).
 TT -> DC (IN REF. 5).
 Q -> T (IN REF. 5).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 G -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).

```
FT CONFLICT 370 370 L -> A (IN REF. 4).
FT CONFLICT 406 409 WPXY -> PYI (IN REF. 5).
FT CONFLICT 448 448 G -> S (IN REF. 5).
FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).
FT HELIX 24 27
FT TURN 28 29
FT STRAND 32 35
FT HELIX 37 40
FT STRAND 42 42
FT TURN 43 44
FT HELIX 53 55
FT STRAND 61 61
FT HELIX 63 67
FT TURN 68 69
FT HELIX 70 74
FT TURN 75 77
FT STRAND 80 83
FT STRAND 87 89
FT STRAND 94 94
FT TURN 95 96
FT STRAND 97 97
FT TURN 100 101
FT STRAND 105 111
FT TURN 113 115
FT TURN 118 119
FT HELIX 120 130
FT TURN 131 132
FT STRAND 134 139
FT STRAND 143 143
FT STRAND 146 146
FT HELIX 150 152
FT TURN 155 156
FT STRAND 159 159
FT HELIX 164 166
FT STRAND 167 167
FT STRAND 172 172
FT HELIX 179 184
FT STRAND 186 188
FT STRAND 193 194
FT STRAND 196 197
FT TURN 199 200
FT HELIX 202 219
FT TURN 220 220
FT STRAND 223 226
FT TURN 227 228
FT HELIX 229 231
FT TURN 234 235
FT HELIX 236 244
FT STRAND 247 250
FT HELIX 257 265
FT TURN 266 266
FT STRAND 269 271
FT HELIX 273 283
FT TURN 286 287
FT HELIX 290 303
FT STRAND 307 309
FT HELIX 311 312
FT TURN 317 318
FT HELIX 322 325
FT STRAND 329 341
FT TURN 345 349
FT STRAND 350 351
FT TURN 352 354
FT HELIX 355 355
FT TURN 361 364
FT HELIX 368 371
FT TURN 372 372

Query Match 100.0%; Score 57; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 SSVDSVFKPF 11

```
Db 151 SSVDSVFKPF 161
|||||
RESULT 3
AMYB_ASPAW STANDARD; PRT; 499 AA.
AC 002906;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha amylase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase B).
GN AMYB.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UVR143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berta R.M.;
RT *Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.;
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X52756; CAA36967.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR00461; Alpha_amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 499
FT ACT_SITE 227 227
FT ACT_SITE 231 231
FT ACT_SITE 251 251
FT ACT_SITE 318 318
FT DISULFID 51 59
FT DISULFID 171 185
FT DISULFID 261 304
FT DISULFID 461 496
FT CARBOHYD 218 218
SQ SEQUENCE 499 AA; 54921 MW; 740B96B1BC01A8A CRC64;
-----
```

```
Query Match 100.0%; Score 57; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 SSVDSVFKPF 11

RESULT 4
AMY_ASPSH

```

ID  AMY ASPSH          STANDARD:          PRT:          499 AA.
AC  P30292;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY.
OS  Aspergillus shirousamli.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC  Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX  NCBI_TaxID=5070;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92323146; PubMed=1368777;
RA  Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT  "Cloning of the alpha-amylase cDNA of Aspergillus shirousamli and its
RT  expression in Saccharomyces cerevisiae."
RL  Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC  -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; D10461; BAA01255.1; -.
DR  PIR; J50663; JS0663.
DR  HSSP; P10529; 7TAA.
DR  InterPro: IPR000461; Alpha-amylase.
DR  Pfam: PF00128; alpha-amylase; 1.
KW  Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW  Glycoprotein; Signal.
FT  CHAIN          1          21
FT  ACT_SITE       22          499  ALPHA-AMYLASE.
FT  ACT_SITE       227         227
FT  ACT_SITE       231         231  BY SIMILARITY.
FT  ACT_SITE       251         251  BY SIMILARITY.
FT  ACT_SITE       318         318  BY SIMILARITY.
FT  DISULFID       51          59  BY SIMILARITY.
FT  DISULFID       171         185  BY SIMILARITY.
FT  DISULFID       261         304  BY SIMILARITY.
FT  DISULFID       461         496  BY SIMILARITY.
FT  CARBOHYD       218         218  N-LINKED (GLCNAC... ) (BY SIMILARITY).
SQ  SEQUENCE      499 AA; 54852 MW; 1FB7AE5DDA01C03F CRC64;

```

```

Query Match          100.0%; Score 57; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY  1 SSVDSYVERPF 11
    |||
DB  151 SSVDSYVERPF 161

```

```

RESULT# 5
AMY1_DEBOC          STANDARD:          PRT:          512 AA.
AC  P19269;
DT  01-NOV-1990 (Rel. 16, Created)
DT  01-NOV-1990 (Rel. 16, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amylase 1 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY1.

```

```

OS  Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX  NCBI_TaxID=27300;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=ATCC 26077; PubMed=1612414;
RC  MEDLINE=92307400; PubMed=1612414;
RA  Park J.C., Bai S., Tai C.Y., Chun S.B.;
RT  "Nucleotide sequence of the extracellular alpha-amylase gene in the
RT  yeast Schwanniomyces occidentalis ATCC 26077."
RL  FEMS Microbiol. Lett. 72:17-23(1992).
CC  -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -1- ENZYME REGULATION: ALPHA-AMYLASE EXPRESSION UNDERLIES CATABOLITE
CC  REPRESSION BY GLUCOSE.
CC  -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; S77586; AAB21151.2; -.
DR  EMBL; X16040; CAA34162.1; -.
DR  EMBL; X62079; CAA43995.1; -.
DR  EMBL; S38381; AAB22383.2; -.
DR  PIR; S06115; S06115.
DR  PIR; S23355; S23355.
DR  HSSP; P10529; 7TAA.
DR  InterPro: IPR000461; Alpha-amylase.
DR  Pfam: PF00128; alpha-amylase; 1.
KW  Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW  Glycoprotein; Signal.
FT  CHAIN          1          25
FT  ACT_SITE       26          512  ALPHA-AMYLASE 1.
FT  ACT_SITE       242         242  BY SIMILARITY.
FT  ACT_SITE       246         246  BY SIMILARITY.
FT  ACT_SITE       333         333  BY SIMILARITY.
FT  CARBOHYD       233         233  N-LINKED (GLCNAC... ) (PROBABLE).
FT  DISULFID       66          74  BY SIMILARITY.
FT  DISULFID       186         200  BY SIMILARITY.
FT  DISULFID       276         319  BY SIMILARITY.
FT  DISULFID       475         510  BY SIMILARITY.
FT  DISULFID       32          32  M -> K (IN STRAINS CCRC 21164 AND ATCC
FT  VARIANT         36          36  S -> G (IN STRAIN CCRC 21164).
FT  VARIANT         73          73  Y -> I (IN STRAIN ATCC 26077).
FT  VARIANT         80          80  N -> S (IN STRAIN CCRC 21164).
FT  VARIANT         350         350  D -> A (IN STRAINS CCRC 21164 AND ATCC

```

FT CARBOHYD	147	147	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	286	286	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	296	296	N-LINKED (GLCNAC. . .)	(POTENTIAL).

FT	355	355	N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD	303	303	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	355	355	N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 975 AA; 109335 MW; 8F570BD9F05B1CB CRC64;

Query Match 68.4%; Score 39; DB 1; Length 975;
Best Local Similarity 77.8%; Pred. No. 8.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVDYSVEK 9
Db 479 SSIDYSAFK 487

RESULT 8
KIT_FELCA STANDARD; PRT; 978 AA.
ID KIT_FELCA
AC Q28889;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mest/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Felis silvestris catus (cat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RA MEDLINE=95140426; PubMed=7530827;
RX Herbst R., Munemitsu S., Ullrich A.;
RT "Oncogenic activation of v-kit involves deletion of a putative
RT tyrosine-substrate interaction site.";
RL Oncogene 10:369-379(1995).
CC -1- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S76596; AAB3207.1; -.
DR HSSP; P11362; IAGW.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR003599; I9.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003600; IG_1ike.
DR InterPro; IPR001824; Receptor_tyr_kin_III.
DR InterPro; IPR001245; Tyr_Pkinase.
DR Pfam; PF00047; Ig_2.
DR Pfam; PF00069; Pkinase; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00410; IG_1ike; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1. Receptor; Transmembrane;
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 978 MAST/STEM CELL GROWTH FACTOR RECEPTOR.
FT DOMAIN 23 521 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 522 544 POTENTIAL.
FT DOMAIN 545 978 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 590 938 PROTEIN KINASE.
FT NP_BIND 596 604 ATP (BY SIMILARITY).
FT BINDING 624 624 ATP (BY SIMILARITY).
FT ACT_SITE 793 793 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 824 824 BY SIMILARITY.
FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 978 AA; 109449 MW; 6D45472E07440B6B CRC64;

Query Match 68.4%; Score 39; DB 1; Length 978;
Best Local Similarity 77.8%; Pred. No. 8.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVDYSVEK 9
Db 477 SSIDYSAFK 485

RESULT 9
AMY2_SCHPO STANDARD; PRT; 513 AA.
ID AMY2_SCHPO
AC 014154;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Probable alpha-amylase C4A8.01 precursor (EC 3.2.1.1) (1,4-alpha-D-
DE glucan glucanhydrolase).
GN SPAC4A8.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```


DR EMBL: Z98762; CAB11471.1; -
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
 KM Calcium; Glycoprotein; Signal
 FT SIGNAL 1 25
 FT CHAIN 26 513
 FT ACT_SITE 226 226 PROBABLE ALPHA-AMYLASE C4A8.01.
 FT ACT_SITE 230 230 BY SIMILARITY.
 FT ACT_SITE 250 250 BY SIMILARITY.
 FT ACT_SITE 318 318 BY SIMILARITY.
 FT DISULFID 52 60 BY SIMILARITY.
 FT DISULFID 171 184 BY SIMILARITY.
 FT DISULFID 260 304 BY SIMILARITY.
 FT DISULFID 454 488 BY SIMILARITY.
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 513 AA: 58715 MW: 455097F4A28C182 CRC64;

Query Match 66.7%; Score 38; DB 1; Length 513;
 Best Local Similarity 40.0%; Pred. No. 6.7;
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 SVDSYVFKPF 11
 ::||::|||

DB 152 NIDYGIYRPF 161

RESULT 10
 AMYL_SACFI STANDARD: PRT; 494 AA.
 ID AMYL_SACFI
 AC P21567;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Alpha-amylase precursor (EC 3.2.1.1).
 GN ALPL.
 OS Saccharomycopsis fibuligera (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
 OX NCBI_TaxID=4944;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87276512; PubMed=3497057;
 RA Itoh T., Yamashita I., Fukui S.;
 RT "Nucleotide sequence of the alpha-amylase gene (ALPL) in the yeast
 Saccharomycopsis fibuligera.";
 RL FEBS Lett. 219:339-342(1987).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 linkages in oligosaccharides and polysaccharides.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO
 KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X05791; CAA29233.1; -
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KM Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT SIGNAL 26

FT CHAIN 27 494
 FT ACT_SITE 233 233 ALPHA-AMYLASE.
 FT ACT_SITE 237 237 BY SIMILARITY.
 FT ACT_SITE 324 324 BY SIMILARITY.
 FT CARBOHYD 224 224 BY SIMILARITY.
 FT DISULFID 57 65 N-LINKED (GLCNAC. . .) (PROBABLE).
 FT DISULFID 177 191 BY SIMILARITY.
 FT DISULFID 267 310 BY SIMILARITY.
 FT DISULFID 462 493 BY SIMILARITY.
 SQ SEQUENCE 494 AA: 54387 MW: 7F7D8FDD6BF58B67 CRC64;

Query Match 64.9%; Score 37; DB 1; Length 494;
 Best Local Similarity 60.0%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SVDSYVFKPF 11
 ::|||::||

DB 158 SIDYSEYRPF 167

RESULT 11
 DRTS_MAIZE STANDARD: PRT; 521 AA.
 ID DRTS_MAIZE
 AC O81395;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-
 [includes: dihydrofolate reductase (EC 1.5.1.3); Thymidylate synthase
 (EC 2.1.1.45)]).
 GN DRTS.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cox K.M., Robertson D., Fites R.C.;
 RT "Cloning of a cDNA encoding the bifunctional dihydrofolate reductase-
 thymidylate synthase (DHFR-*ts*) in Zea mays.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADP(+) = 7,8-
 dihydrofolate + NADPH.
 CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP =
 dihydrofolate + dTMP.
 CC -1- PATHWAY: ESSENTIAL STEP FOR DE NOVO GLYCINE AND PURINE SYNTHESIS,
 DNA PRECURSOR SYNTHESIS, AND FOR THE CONVERSION OF DUMP TO dTMP.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 DIHYDROFOLATE REDUCTASE FAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE THYMIDYLATE
 SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF073488; AAC26003.1; -
 DR HSSP: P13100; 1C17.
 DR Mendel: 32032; Zeama: 2610; 32032.
 DR InterPro: IPR001796; DHFR.
 DR InterPro: IPR000398; Thymidylat_synth.
 DR Pfam: PF00186; Dihfolate_red; 1.
 DR Pfam: PF00303; thymidylat_synth; 1.
 DR PRINTS: PR00070; DHFR.
 DR PRINTS: PR00108; THYMSNTHASE.
 DR ProDom: PD001180; Thymidylat_synth; 1.
 DR PROSITE: PS00075; DHFR; 1.

DR PROSITE: PS00091: THYMIDYLATE SYNTHASE; 1.
 KW Multifunctional enzyme; Oxidoreductase; Transferase; NADs;
 KM Methyltransferase; Nucleotide biosynthesis; One-carbon metabolism.
 FT DOMAIN 1 196 DIHYDROFOLATE REDUCTASE.
 FT DOMAIN 197 521 THYMIDYLATE SYNTHASE.
 FT ACT SITE 403 403 BY SIMILARITY.
 SQ SEQUENCE 521 AA; 58965 MW; 81266f852625f06 CRC64;

Query Match 63.2%; Score 36; DB 1; Length 521;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDYSVFKPF 11
 Db 166 VDYSVFKPF 174

RESULT 12
 ID DPO3_THEMEA STANDARD; PRT; 1367 AA.
 AC Q9ZHF6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA polymerase III polC-type (EC 2.7.7.7) (PolIII).
 GN POLC OR TM0576.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99045593; PubMed=9826752;
 RA Huang Y.P., Ito J.;
 RT "The hyperthermophilic bacterium Thermotoga maritima has two different
 RT classes of family C DNA polymerases: evolutionary implications.";
 RL Nucleic Acids Res. 26:5300-5309(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RA MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Hatt D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Uitterlind T.R., Malek J.A., Linher C.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima.";
 RL Nature 399:323-329(1999).
 CC -1- FUNCTION: REQUIRED FOR REPLICATIVE DNA SYNTHESIS. THIS DNA
 CC POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
 CC + [DNA] (N).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. POLC
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF065313; AAC80438.1; -;
 CC EMBL: AE001732; AAD35661.1; -;
 CC TIGR: TM0576; -;
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR004013; PHP_C.
 DR InterPro: IPR003141; PHP_N.

DR InterPro: IPR002309; tRNA-synt-2.
 DR Pfam: PF00929; Exonuclease; 1.
 DR Pfam: PF02811; PHP_C; 1.
 DR Pfam: PF02231; PHP_N; 1.
 DR Pfam: PF01336; tRNA_ant1; 1.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00479; EXOIII; 1.
 DR SMART: SM00481; POLIITAC; 1.
 KW Transferase; DNA-directed DNA polymerase; DNA replication; Hydrolyase;
 KW Nuclease; Exonuclease; Complete proteome.
 FT DOMAIN 1 358 521 EXONUCLEASE.
 SQ SEQUENCE 1367 AA; 155362 MW; EE5910FA70591F84 CRC64;

Query Match 63.2%; Score 36; DB 1; Length 1367;
 Best Local Similarity 50.0%; Pred. No. 47;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 SVDYSVFKPF 11
 Db 537 TIDYALKPF 546

RESULT 13
 ID IF4E_APLCA STANDARD; PRT; 215 AA.
 AC 077210;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Eukaryotic translation Initiation factor 4E (eIF4E) (mRNA
 DE cap-binding protein) (eIF-4F 25 kDa subunit).
 OS Aplysia californica (California sea hare).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anaspidaea;
 OX Aplysiidae; Aplysia.
 RN NCBI_TaxID=6500;
 RN [1]
 RP SEQUENCE FROM N.A., AND PHOSPHORYLATION.
 RC TISSUE=Nerve;
 RA MEDLINE=99009053; PubMed=9792652;
 RA Dyer J.R., Peipo A.M., Yanow S.K., Sossin W.S.;
 RT "Phosphorylation of eIF4E at a conserved serine in Aplysia.";
 RL J. Biol. Chem. 273:29469-29474(1998).
 CC -1- FUNCTION: RECOGNIZES AND BINDS THE 7-METHYLGUANOSINE-CONTAINING
 CC MRNA "CAP" DURING AN EARLY STEP IN THE INITIATION OF PROTEIN
 CC SYNTHESIS AND FACILITATES RIBOSOME BINDING BY INDUCING THE
 CC UNWINDING OF THE MRNA SECONDARY STRUCTURES.
 CC -1- SUBUNIT: EIF4E IS A TRIMER COMPOSED OF EIF4E, EIF4G AND EIF4A
 CC (WHICH CAN CYCLE IN AND OUT OF THE COMPLEX).
 CC -1- PTM: PHOSPHORYLATION INCREASE THE ABILITY OF THE PROTEIN TO BIND
 CC TO MRNA CAPS AND TO FORM THE EIF4E COMPLEX.
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF085810; AAC36720.1; -;
 CC HSSP: P07260; IAP8.
 DR InterPro: IPR001040; eIF_4E.
 DR Pfam: PF01652; IF4E; 1.
 DR PRODOM: PD003697; eIF_4E; 1.
 DR PROSITE: PS00813; IF4E; 1.
 KW Initiation factor; Protein biosynthesis; RNA-binding; Phosphorylation.
 FT MOD_RES 207 207 PHOSPHORYLATION (BY PKC).
 SQ SEQUENCE 215 AA; 24646 MW; 9EE00CB6DE8162E7 CRC64;

Query Match 61.4%; Score 35; DB 1; Length 215;

Best Local Similarity 77.8%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSVDSVFK 9
| | | | |
DB 85 SCDISVFK 93

RESULT 14
ARCA_BORAF STANDARD; PRT: 409 AA.
ID ARCA_BORAF
AC 051896; Created
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Arginine deiminase (EC 3.5.3.6) (ADI) (Arginine dihydrolase) (AD).
GN ARCA.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
CX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R-1p3;
RX MEDLINE=98065594; PubMed=9402027;
RA Casjens S., Murphy M., Delange M., Sampson L., van Vugt R.,
Huang W.M.;
RT "Telomers of the linear chromosomes of Lyme disease spirochaetes:
nucleotide sequence and possible exchange with linear plasmid
telomers.";
RL Mol. Microbiol. 26:581-596(1997).
CC -I- CATALYTIC ACTIVITY: L-arginine + H(2)O = L-citrulline + NH(3).
CC -I- PATHWAY: Arginine degradation via arginine deiminase; first step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -I- SIMILARITY: BELONGS TO THE ARGININE DEIMINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch)

CC
CC EMBL: AF008219; AAB93998.1; -
DR InterPro: IPR003876; Arg_deiminase.
DR Pfam: PF02726; Arg_deiminase; 1.
DR PRINTS: PR01466; ARGDEIMINASE.
KW Hydrolase; Arginine metabolism.
SQ SEQUENCE 409 AA; 4681 MW; 0524047ECA16649C CRC64;

Query Match 61.4%; Score 35; DB 1; Length 409;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
: | | | | |
DB 278 TQIDSVFTSF 288

RESULT 15
ARCA_BORBU STANDARD; PRT: 410 AA.
ID ARCA_BORBU
AC 051781;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Arginine deiminase (EC 3.5.3.6) (ADI) (Arginine dihydrolase) (AD).
GN ARCA OR BB0841.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=980655943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Ginn M.,
Dougenberry B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Matthey L., McDonald L., Artach P., Bowman C.,
Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi.";
RL Nature 390:580-586(1997).
CC -I- CATALYTIC ACTIVITY: L-arginine + H(2)O = L-citrulline + NH(3).
CC -I- PATHWAY: Arginine degradation via arginine deiminase; first step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -I- SIMILARITY: BELONGS TO THE ARGININE DEIMINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch)

CC
CC EMBL: AE001183; AAC67191.1; -
DR TIGR: BB0841; -
DR InterPro: IPR003876; Arg_deiminase.
DR Pfam: PF02726; Arg_deiminase; 1.
DR PRINTS: PR01466; ARGDEIMINASE.
KW Hydrolase; Arginine metabolism; Complete proteome.
SQ SEQUENCE 410 AA; 46843 MW; C2DF4F58E5848PAB CRC64;

Query Match 61.4%; Score 35; DB 1; Length 410;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
: | | | | |
DB 279 TQIDSVFTSF 289

Search completed: June 13, 2002, 09:11:51
Job time: 1031 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:55:49 ; Search time 73.31 Seconds
(without alignments)
14.418 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160

Perfect score: 57

Sequence: 1 SSVDSYVFKPF 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

PIR_71:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	57	100.0	478 2 JK0201	alpha-amylose (EC
2	57	100.0	498 2 A48305	alpha-amylose (EC
3	57	100.0	499 1 A4831	alpha-amylose (EC
4	57	100.0	499 1 ALA51	alpha-amylose (EC
5	57	100.0	499 2 B48305	alpha-amylose (EC
6	57	100.0	499 2 JS0663	alpha-amylose (EC
7	57	100.0	499 2 JT0466	alpha-amylose (EC
8	57	100.0	499 2 JT0588	alpha-amylose (EC
9	45	78.9	512 2 S06115	alpha-amylose (EC
10	45	78.9	512 2 S23355	alpha-amylose (EC
11	44	77.2	484 1 A35282	alpha-amylose (EC
12	40	70.2	489 2 AF3457	D-hydantoinase (EC
13	39	68.4	4116 2 T13719	calo protein - fru
14	38	66.7	425 2 H82183	conserved hypotet
15	38	66.7	430 2 T37549	hypothetical prote
16	38	66.7	513 2 T38770	alpha-amylose a pr
17	38	66.7	1869 2 A59290	class V chitin syn
18	37	64.9	494 1 ALBYAF	alpha-amylose (EC
19	36	63.2	112 2 T47522	hypothetical prote
20	36	63.2	173 2 T27650	hypothetical prote
21	36	63.2	300 2 G86765	hypothetical prote
22	36	63.2	353 2 T19522	hypothetical prote
23	36	63.2	409 2 T46063	hypothetical prote
24	36	63.2	491 2 T38448	probable alpha-amy
25	36	63.2	521 2 T01684	dihydroxylate redu
26	36	63.2	541 2 T43862	FAD flavoprotein o
27	36	63.2	624 2 T39102	hypothetical prote
28	36	63.2	1367 2 C72360	DNA polymerase III
29	36	63.2	2692 2 T23768	hypothetical prote

30	35	61.4	116 2 H98221	hypothetical prote
31	35	61.4	153 2 AH3064	hypothetical prote
32	35	61.4	363 2 T44564	probable thiamin b
33	35	61.4	377 2 T50563	beta-1,3-glucanase
34	35	61.4	410 2 A29393	hemocyanin beta-c
35	35	61.4	410 2 H70204	arginine deiminase
36	35	61.4	492 2 F86384	probable protein f
37	35	61.4	497 2 T52308	very-long-chain fa
38	35	61.4	562 2 T34319	hypothetical prote
39	35	61.4	1498 2 S78102	chitin synthase (E
40	35	61.4	1852 2 JC5546	chitin synthase (E
41	34	59.6	201 2 AH2231	hypothetical prote
42	34	59.6	213 2 H83839	hypothetical prote
43	34	59.6	370 2 C71052	hypothetical prote
44	34	59.6	407 2 T30469	hypothetical prote
45	34	59.6	418 2 E96687	hypothetical prote

ALIGNMENTS

RESULT 1

JK0201

alpha-amylose (EC 3.2.1.1) - Aspergillus oryzae

N:Alternate names: glycogenase; Taka-amylose A

C:Species: Aspergillus oryzae

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998

C:Accession: JK0201

R: Toda, H.; Kondo, K.; Narita, K.

Proc. Jpn. Acad. 58B, 208-212, 1982

A:Title: The complete amino acid sequence of Taka-amylose A.

A:Reference number: JK0201

A:Molecule type: protein

A:Residues: 1-478 <100>

C:Comment: One atom of calcium per molecule is essential for the activity.

C:Comment: This enzyme is a glycoprotein.

C:Comment: See also PIR:JT0466 and PIR:JS0240.

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:173-300/Domains: alpha-amylose core homology <AMY>

F:197/Binding site: carboxylate (Asn) (covalent) #status experimental

F:210/230/297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 57; DB 2; Length 478;

Best local similarity 100.0%; Pred. No. 0.0049;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVDSYVFKPF 11

DB 130 SSVDSYVFKPF 140

RESULT 2

A48305

alpha-amylose (EC 3.2.1.1) A precursor - Aspergillus awamori

C:Species: Aspergillus awamori

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997

C:Accession: A48305

R: Kornman, D.R.; Bayliss, F.T.; Barnett, C.C.; Cammona, C.L.; Kodama, K.H.; Royer, T.J

Curr. Genet. 17, 203-212, 1990

A:Title: Cloning, characterization, and expression of two alpha-amylose genes from As

A:Reference number: A48305; MUID:90254827

A:Accession: A48305

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-498 <KOR>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A.Pathway: glycogen/starch degradation
 C.Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C.Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F.194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 57; DB 2; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SSVDSVFKPF 11
 |||||
 Db 151 SSVDSVFKPF 161

RESULT 3
 ALAS1
 alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
 N.Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
 C.Species: Aspergillus oryzae
 C.Date: 30-Sep-1991 #sequence-revision 30-Sep-1991 #text-change 15-Sep-2000
 C.Accession: S04548; A33214; J50240; A91930; A93767; A10627
 R.Witzel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A.Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon
 A.Reference number: S04548; MUID:89237897
 A.Accession: S04548
 A.Molecule type: DNA
 A.Residues: 1-499 <WIR>
 A.Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
 A.Genetics: AMY1
 A.Accession: A33214
 A.Molecule type: mRNA
 A.Residues: 1-499 <MID>
 A.Cross-references: GB:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
 R.Genes, M.J.; Dove, M.J.; Selligy, V.L.
 Gene 79, 107-117, 1989
 A.Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
 A.Reference number: J50240; MUID:89378767
 A.Accession: J50240
 A.Molecule type: DNA
 A.Residues: 1-499 <GEN>
 A.Genetics: AMY2
 A.Note: the authors refer to this as isozyme II
 R.Isemura, S.; Ikenaka, T.
 J. Biochem. 74, 1-10, 1973
 A.Reference number: A91930; MUID:74001521
 A.Accession: A91930
 A.Molecule type: Protein
 A.Residues: 206-225 <ISE>
 R.Narita, K.
 Proc. Jpn. Acad. 51, 285-290, 1975
 A.Reference number: A93767
 A.Accession: A93767
 A.Molecule type: Protein
 A.Residues: 434-443,446-447, 'Q',449-458, 'GTW',459-464,467-468, 'B',470, 'B',472-499 <NMR>
 J. Blochem. 95, 697-702, 1984
 A.Title: Structure and possible catalytic residues of Taka-amylase A.
 A.Reference number: A37454; MUID:8421370
 A.Contents: annotation: X-ray crystallography, 3.0 angstroms
 R.Swift, H.J.; Brady, L.; Dewende, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.
 submitted to the Brookhaven Protein Data Bank, August 1992
 A.Reference number: A51548; PDB:6TAA
 A.Contents: annotation: X-ray crystallography, 2.1 angstroms, residues 22-497
 C.Comment: One atom of calcium per molecule is essential for activity.
 C.Genetics: <AMW1>
 A.Gene: amy1
 A.Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C.Genetics: <AMW2>
 A.Gene: amy2; AmyII
 A.Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C.Function:

A.Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A.Pathway: glycogen/starch degradation
 C.Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C.Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po
 F.1-21/Domain: signal sequence #status predicted <SIG>
 F.22-499/Product: alpha-amylase 1 #status experimental <MAT>
 F.194-321/Domain: alpha-amylase core homology <AMY>
 F.151-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F.218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F.227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 57; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SSVDSVFKPF 11
 |||||
 Db 151 SSVDSVFKPF 161

RESULT 4
 ALAS3
 alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
 N.Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A
 C.Species: Aspergillus oryzae
 C.Date: 30-Sep-1991 #sequence-revision 30-Sep-1991 #text-change 18-Jun-1999
 C.Accession: S04549; A33215; A44713
 R.Witzel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A.Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
 A.Reference number: S04548; MUID:89237897
 A.Accession: S04549
 A.Molecule type: DNA
 A.Residues: 1-499 <WIR>
 A.Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 A.Accession: A33215
 A.Status: translation not shown
 A.Molecule type: mRNA
 A.Residues: 1-499 <MID>
 A.Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 R.Genes, M.J.; Dove, M.J.; Selligy, V.L.
 Gene 79, 107-117, 1989
 A.Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each contain
 A.Reference number: J50240; MUID:89378767
 A.Accession: A44713
 A.Molecule type: DNA
 A.Residues: 1-499 <GEN>
 A.Note: the authors refer to this as isozyme I
 R.Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A.Title: Structure and possible catalytic residues of Taka-amylase A.
 A.Reference number: A37454; MUID:8421370
 A.Contents: annotation: X-ray crystallography, 3.0 angstroms
 C.Comment: One atom of calcium per molecule is essential for activity.
 C.Genetics:
 A.Gene: amy3; AmyI
 A.Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C.Function:
 A.Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A.Pathway: glycogen/starch degradation
 C.Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C.Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
 F.1-21/Domain: signal sequence #status predicted <SIG>
 F.22-499/Product: alpha-amylase 3 #status experimental <MAT>
 F.194-321/Domain: alpha-amylase core homology <AMY>
 F.51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F.142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F.218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F.231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 57; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
DB 151 SSVDSYVFKPF 161

RESULT 5

alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
B48305
C:Species: Aspergillus awamori
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: B48305
R:Korman, D. R.; Bayliss, F. T.; Barnett, C. C.; Carmona, C. L.; Kodama, K. H.; Royer, T. J.;
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus
A:Reference number: M48305; MUID:90254827
A:Accession: B48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 57; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
DB 151 SSVDSYVFKPF 161

RESULT 6

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
J50663
C:Species: Aspergillus sp.
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
C:Accession: J50663
R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Biotech. Biotechnol. Biochem. 56, 174-179, 1992
A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its expression
A:Reference number: J50663; MUID:92323146
A:Accession: J50663
A:Molecule type: mRNA
A:Residues: 1-499 <SHI>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMP>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 57; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
DB 151 SSVDSYVFKPF 161

Query Match 100.0%; Score 57; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
DB 151 SSVDSYVFKPF 161

RESULT 7

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
JT0466
N:Alternate names: glycosidase; Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JT0466
R:Tada, S.; Jimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus
A:Reference number: JT0466
A:Accession: JT0466
A:Molecule type: DNA
A:Residues: 1-499 <TAD>
C:Comment: See also PIR:JK0201 and PIR:J50240.
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:221,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 57; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
DB 151 SSVDSYVFKPF 161

RESULT 8

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
JN0588
N:Alternate names: Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
C:Accession: JN0588
R:Tukugoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsuboi, A.; Uchida, S.
Gene 84, 319-327, 1989
A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
A:Reference number: JN0588; MUID:90128276
A:Accession: JN0588
A:Molecule type: mRNA
A:Residues: 1-499 <TSU>
C:Comment: The alpha amylases are encoded by multigene family.
C:Genetics:
A:Gene: Taa-G1
A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 57; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||

Db 151 SSVDSVFKPF 161

RESULT 9

S06115
alpha-amylase (EC 3.2.1.1) AMY1 precursor - yeast (Schwanniomyces occidentalis)
C:Species: Schwanniomyces occidentalis
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 22-Jun-1999
C:Accession: S06115
R:Strasser, A.W.M.; Selk, R.; Dohmen, R.J.; Niermann, T.; Bielefeld, M.; Seeboth, F.; Tu
Eur. J. Biochem. 184, 699-706, 1999
A:Title: Analysis of the alpha-amylase gene of Schwanniomyces occidentalis and the secret
A:Reference number: S06115; MUID:90032653
A:Accession: S06115
A:Molecule type: DNA
A:Residues: 1-512 <STR>
A:Cross-references: EMBL:X16040; NID:g4882; PIDN:CNA34162.1; PID:g4883
C:Genetics:
A:Gene: AMY1
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-25/Domain: signal sequence #status predicted <Sig>
F:26-512/Product: alpha-amylase #status predicted <MAT>
F:209-336/Domain: alpha-amylase core homology <AMY>

Query Match 78.9%; Score 45; DB 2; Length 512;

Best Local Similarity 81.8%; Pred. No. 1.1;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDSVFKPF 11
Db 166 SSVDSVFKPF 176

RESULT 10

S23355
alpha-amylase (EC 3.2.1.1) precursor - yeast (Schwanniomyces occidentalis)
C:Species: Schwanniomyces occidentalis
C>Date: 22-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999
C:Accession: S23355
R:Wu, F.M.; Wang, T.T.; Hsu, W.H.
FEMS Microbiol. Lett. 82, 313-318, 1991
A:Title: The nucleotide sequence of Schwanniomyces occidentalis alpha-amylase gene.
A:Reference number: S23355
A:Accession: S23355
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <WU>
A:Cross-references: EMBL:X62079; NID:g4880; PIDN:CAA43995.1; PID:g4881
A>Note: the authors translated the codon AGA for residue 21 as Pro, CTT for residue 61
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase
F:209-336/Domain: alpha-amylase core homology <AMY>

Query Match 78.9%; Score 45; DB 2; Length 512;

Best Local Similarity 81.8%; Pred. No. 1.1;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDSVFKPF 11
Db 166 SSVDSVFKPF 176

RESULT 11

A35282
alpha-amylase (EC 3.2.1.1) - Aspergillus niger
C:Species: Aspergillus niger

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A35282

R:Boel, E.; Brady, L.; Brzozowski, A.M.; Derewenda, Z.; Dodson, G.G.; Jensen, V.J.; P

Biochemistry 29, 6244-6249, 1990

A:Title: Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-angstrom

A:Reference number: A35282; MUID:91002514

A:Accession: A35282

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-484 <BOE>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:173-300/Domain: alpha-amylase core homology <AMY>

Query Match 77.2%; Score 44; DB 1; Length 484;

Best Local Similarity 72.7%; Pred. No. 1.6;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDSVFKPF 11
Db 130 NDVDSVFKPF 140

RESULT 12

AF3457
D-hydantoinase (EC 3.5.2.2) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AF3457
R:Delvecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3457
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-489 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52825.1; PID:g17983665; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11644
A:Map position: 1
C:Superfamily: allantoinase; Bacillus dihydroorotase homology
C:Keywords: hydrolase

Query Match 70.2%; Score 40; DB 2; Length 489;

Best Local Similarity 63.6%; Pred. No. 9.2;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVDSVFKPF 11
Db 412 SSIDYVFEKF 422

RESULT 13

T13719
calo protein - fruit fly (Drosophila melanogaster) (fragment)

C:Species: Drosophila melanogaster

C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C:Accession: T13719

R:Xu, X.Z.; Wes, P.D.; Chen, H.; Li, H.S.; Yu, M.; Morgan, S.; Liu, Y.; Montell, C.

J. Biol. Chem. 273, 31297, 1998

A:Title: Retinal targets for calmodulin include proteins implicated in synaptic trans

A:Reference number: Z17709; MUID:99030403

A:Accession: T13719

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-4116 <XUX>
A:Cross-references: EMBL:Y17920; NID:g3893108; PIDN:CAA76940.1; PID:g3893109
C:Genetics:
A:Gene: calo
A:Cross-references: Flybase:Flyg0011230

Query Match 68.4%; Score 39; DB 2; Length 4116;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 DYSVFKPF 11
11:111111
DB 4013 DYTIVKPF 4020

RESULT 14

H82183
conserved hypothetical protein VC1567 [imported] - *Vibrio cholerae* (strain N16961 serog
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: H82183
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; M01D:20406833
A:Accession: H82183
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <HEI>
A:Cross-references: GB:AE004234; GB:AE003852; NID:g9656070; PIDN:AAF94721.1; GSPDB:GN001
C:Genetics:
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
A:Gene: VC1567
A:Map position: 1

Query Match 66.7%; Score 38; DB 2; Length 425;
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SVDYSVFKPF 11
1:111111
DB 209 SPIDSGFTPF 219

RESULT 15

T37549

hypothetical protein SPAC1167.06c - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37549
R:McLean, J.; Harris, D.; Bartell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL data library, September 1997
A:Reference number: Z21722
A:Accession: T37549
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-430 <MCL>
A:Cross-references: EMBL:Z99161; PIDN:CAB16211.1; GSPDB:GN00066; SPDB:SPAC1167.06c
C:Genetics:
A:Experimental source: strain 972h-; cosmid c1167
A:Gene: SPDB:SPAC1167.06c
A:Map position: 1

Query Match 66.7%; Score 38; DB 2; Length 430;
Best Local Similarity 70.0%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 SVDYSVFKPF 11

DB 21 SVDYSQFKAF 30
1:1111111

Search completed: June 13, 2002, 08:55:51
Job time: 371 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:53:26 ; Search time 158.68 Seconds
(without alignments)
4.900 Million cell updates/sec

Title: US-09-710-339-2_COPY_161_167

Perfect score: 40

Sequence: 1 SSODYFH 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	478	15 AAR46065	Mutant alpha-amylase
2	40	100.0	478	16 AAR79025	Mature taka-amylase
3	40	100.0	478	16 AAR72450	Aspergillus oryzae
4	40	100.0	478	16 AAR78270	Aspergillus oryzae
5	40	100.0	478	16 AAM14500	Aspergillus oryzae
6	40	100.0	498	22 AAB84206	Amino acid sequence
7	37	92.5	493	17 AAR88212	Alpha-amylase. Th
8	34	85.0	52	22 AAC008447	Human polypeptide
9	34	85.0	1222	21 AAG36453	Arabidopsis thaliana
10	34	85.0	1257	21 AAG36452	Arabidopsis thaliana
11	34	85.0	1275	21 AAG36451	Arabidopsis thaliana

12	33	82.5	124	22 AAO05772	Human polypeptide
13	32	80.0	501	22 ABG15126	Novel human diaph
14	32	80.0	529	22 ABB70891	Drosophila melanog
15	32	80.0	583	18 AAR26654	Human PANC1A, assoc
16	32	80.0	594	21 AAB56501	Human prostate can
17	32	80.0	700	20 AAT33673	B. bassiana POPS r
18	32	80.0	715	22 AAT79737	Human protein SEQ
19	32	80.0	1118	19 AAM82395	Human ubp protein
20	32	80.0	1128	22 ABG11843	Novel human diaph
21	32	80.0	1130	22 ABG22077	Human human diaph
22	31	77.5	69	20 AAM67870	Human secreted pro
23	31	77.5	89	22 AAO06897	Human polypeptide
24	31	77.5	104	21 AAY90470	Yeast Sm11 (suppre
25	31	77.5	119	21 AAT31764	Novel human secret
26	31	77.5	123	22 AAO08298	Human polypeptide
27	31	77.5	191	22 AAM80917	Human haematologic
28	31	77.5	191	22 AAM81318	Human haematologic
29	31	77.5	191	22 AAM81390	Human haematologic
30	31	77.5	191	22 AAM81634	Human haematologic
31	31	77.5	191	22 AAM81672	Human haematologic
32	31	77.5	208	22 ABB59074	Drosophila melanog
33	31	77.5	275	22 AAU08723	Human clone h01143
34	31	77.5	275	22 AAB73980	Human starazin-11
35	31	77.5	343	21 AAB43994	Human cancer assoc
36	31	77.5	374	21 AAB42328	Human ORFX ORF2092
37	31	77.5	461	20 AAY35506	Chlamydia pneumoni
38	31	77.5	474	21 AAY56509	Human Jurkat cell
39	31	77.5	500	22 AAE10164	RNA helicase conse
40	31	77.5	614	22 ABB44571	Mouse wound healin
41	31	77.5	614	22 ABB44572	Human wound healin
42	31	77.5	615	22 ABB44573	Mouse wound healin
43	31	77.5	2427	21 AAY51843	Marine embryo macr
44	30	75.0	25	17 AAM01178	Serine protease Pf
45	30	75.0	25	22 AAB50574	Flea serine protea

ALIGNMENTS

ID	RESULT	1	ALIGNMENTS
AA46065	AA46065 standard; protein; 478 AA.		
AC	AA46065;		
DT	18-JUL-1994 (first entry)		
XX			
DE	Mutant alpha-amylase.		
KW	Methionine substitution; stability; activity; detergent;		
KW	dishwashing agents; liquidation agents.		
XX			
OS	Aspergillus oryzae.		
XX			
PN	W09402597-A.		
XX			
PD	03-FEB-1994.		
XX			
PF	06-JUL-1993; 93WO-DK00230.		
XX			
PR	23-JUL-1992; 92DK-0000946.		
PR	16-DEC-1992; 92DK-0001503.		
PR	15-MAR-1993; 93DK-0000292.		
XX			
PA	(NOVO) NOVO-NORDISK AS.		
XX			
PI	Bisgard-frantzen H, Svendsen A;		
XX			
DR	WPI; 1994-048855/06.		
PT	Mutant alpha-amylase from Bacillus species comprising a		
PT	methionine substitution - with improved stability and activity at		
PT	low pH, for use in detergents, dishwashing agents and		

PF Liquefaction agents
XX
PS Claim 1; Page 7; 20pp; English.
XX
CC The sequence os that of the *Aspergillus oryzae* alpha amylase, sold
CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC be mutated by substitution of one or more of its methionine residues
CC for any amino acid other than cysteine. The mutant alpha-amylase
CC exhibits a better activity level and better stability in the
CC presence of oxidising agents than previous mutant alpha amylases,
CC and improved thermostability at moderately low pH. The enzyme can
CC be used as an additive for detergents, dishwashing agents and
CC liquifaction agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 40; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|111111|
Db 141 ssqdyfh 147

RESULT 2
AAR79025 100.0%; Score 40; DB 15; Length 478;
ID AAR79025 standard; protein: 478 AA.
XX
AC AAR79025;
XX
DT 22-MAR-1996 (first entry)
XX
DE Mature taka-amylase A.
XX
KW Wild type; neopullulanase; B. stearotherophilus; mutant;
KW food industry; modification; hydrophobicity; replacement
KW insertion; deletion.
XX
OS *Aspergillus oryzae*.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
PN JP07177891-A.
XX
PD 18-JUL-1995.
XX
PF 31-OCT-1994; 94JP-0288658.
XX
PR 12-NOV-1993; 93JP-0306096.
XX
PA (EZAK) EZAKI GLICO CO.
PA (NIDE) NEC CORP.
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site - increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
PS Disclosure: Page 10-11; 18pp; English.
XX
CC This sequence represents the mature form of taka-amylase from *A. oryzae*.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the

CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC hydrophobic amino acid from the selected site. The method was used
CC to produce neopullulanases Y377F, S422V and M375L.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 40; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|111111|
Db 141 ssqdyfh 147

RESULT 3
AAR72450 100.0%; Score 40; DB 16; Length 478;
ID AAR72450 standard; protein: 478 AA.
XX
AC AAR72450;
XX
DT 01-DEC-1995 (first entry)
XX
DE *Aspergillus oryzae* alpha amylase (mature protein).
XX
KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KW washing; *Bacillus licheniformis*; *Bacillus amyloliquefaciens*;
KW *Bacillus stearotherophilus*; dyeing; bleaching; scouring; textile;
KW thermostable.
XX
OS *Aspergillus oryzae*.
XX
PN W09510603-A.
XX
PD 20-APR-1995.
XX
PF 05-OCT-1994; 94WO-DK00370.
XX
PR 08-OCT-1993; 93DK-0001133.
PR 02-FEB-1994; 94DK-0000140.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Svendsen A, Thel Jensen M;
PI Van der zee P, Bisgardfrantzen H, Borchert T;
XX
DR WPI; 1995-161790/21.
XX
PT New *Bacillus* derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance
XX
PS Disclosure: Page 75-76; 105pp; English.
XX
CC Variant alpha amylase enzymes which have improved washing and/or
CC as detergent additives. The enzymes have one or more amino acid
CC residues added, deleted or substituted. The variants can also be
CC used for textile desizing prior to scouring, bleaching and dyeing.
CC The variants have improved thermostability, acid/alkaline stability;
CC low temperature optimum; pH optimum; higher hydrolysis velocity and
CC improved tolerance to other composition constituents, e.g. oxidation
CC agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 40; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
|111111|

FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to this
FT fragment is deleted or replaced with a fragment
FT corresponding to 102-199 of AA14498; claim 42"

Misc-difference 165..177

/label= loop 3 modification region

"at least one amino acid residue of a parent
alpha-amylase (used as a template for a variant)
corresponding to 195-202 of AA14499 is deleted
or replaced with a fragment corresponding to
this fragment; claim 21"

Misc-difference 166..173

/label= loop 3 modification region

"preferred region where at least one amino acid
residue of a parent alpha-amylase (used as a
template for a variant) corresponding to 196-198
of AA14499 is deleted or replaced with a
fragment corresponding to this fragment; claim
23" Misc-difference 181..184

/note= "an amino acid fragment corresponding to this
region is deleted from the parent sequence of a
variant Fungamyl; claim 43"

Misc-difference 291..313

/label= loop 8 modification region

"at least one amino acid residue of a parent
alpha-amylase (used as a template for a variant)
corresponding to 322-346 of AA14498 is deleted
or replaced with a fragment corresponding to
this fragment; claim 36"

Misc-difference 297..313

/label= loop 8 modification region

"preferred region where at least one amino acid
residue of a parent alpha-amylase (used as a
template for a variant) corresponding to 323-345
of AA14498 is deleted or replaced with a
fragment corresponding to this fragment; claim
38"

MO9623874-A1.

08-AUG-1996.

96WO-DK00057.

05-FEB-1996; 95DK-0001256.
PR 10-NOV-1995; 95DK-0000128.
PR 03-FEB-1995; 95DK-0000128.
PR 23-OCT-1995; 95DK-0001192.

(NOVO) NOVO-NORDISK AS.

Bisgard-frantzen H, Borchert TV, Svendsen A;
WPI; 1996-371424/37.

Alpha-amylase variants and methods of production - have altered
stability

Disclosure; Page 87-88; 171pp: English.

The present sequence is the mature Aspergillus oryzae alpha-amylase (A.).
Variants of parent termamyl- and fungamyl-like alpha-amylases (and
methods of constructing them) are claimed. Examples of variants are
featured above. The variants have altered properties such as calcium
dependency, substrate binding and stability. Also one or more proline or
cysteine residues in the variant is modified or replaced with a
non-proline or non-cysteine residue such as alanine. The variants can be
used for (dis)washing, as detergent additives or for fabric desizing on
starch liquefaction. They can also be used for the production of
sweeteners and ethanol from starch. See also AA14498-99.

Sequence 478 AA;

Query Match	100.0%;	Score 40;	DB 17;	Length 478;
Best Local Similarity	100.0%;	Pred. No. 6.7;		
Matches 7;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1 SSQDYFH 7			
DB	141 ssqdyfh 147			
RESULT 6				
ID	AAB84206	standard; Protein; 498 AA.		
XX	AAB84206			
XX	AAB84206;			
XX	06-AUG-2001	(first entry)		
XX				
DE	Amino acid sequence of a fungamyl-like alpha-amylase.			
XX				
KW	Fungamyl-like alpha-amylase; glucanase; dextrinisation; maltose;			
KW	alcohol; starch; dough improver; brewing; starch liquification.			
XX				
OS	Aspergillus oryzae.			
XX				
PN	WO200134784-A1.			
XX				
PD	17-MAY-2001.			
XX				
XX				
FP	10-NOV-2000; 2000WO-DK00626.			
XX				
PR	10-NOV-1999; 99DK-0001617.			
XX				
XX				
PA	(NOVO) NOVOZYMES AS.			
XX				
PI	Bisgard-Frantzen H, Svendsen A, Pedersen S;			
XX				
DR	WPI; 2001-367478/38.			
XX				
DR	N-PSDB; AAF90208.			
XX				
PT	New variant of Fungamyl-like alpha-amylase, useful for production of			
PT	maltose syrups. Includes mutations that improve stability against heat			
PT	and acidic pH -			
XX				
XX				
PS	Claim 1; Page 42-45; 49pp; English.			
XX				
CC	The present sequence represents a fungamyl-like alpha-amylase. The			
CC	specification describes variants of this fungamyl-like alpha-amylase,			
CC	which have an alteration in one the amino acid regions 98-110, 150-160,			
CC	161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion			
CC	or substitution of an amino acid or an insertion of an amino acid			
CC	downstream of a particular position. The variants retain alpha-amylase			
CC	activity, and have better heat stability and/or stability at acidic pH,			
CC	relative to wild-type enzyme. The variants can therefore be used at			
CC	higher temperatures (more efficient conversion or faster reaction), and			
CC	have reduced need for cooling and reduced risk of contamination). The			
CC	variants may also be used in conjunction with other enzymes,			
CC	particularly glucanase during dextrinisation. The variants are			
CC	used to produce syrups, particularly of high maltose content, or alcohol,			
CC	from starch, as dough improver for baked goods; in brewing, to increase			
CC	fermentability of the wort; and for liquefaction of starch.			
XX				
SO	Sequence 498 AA;			
Query Match 100.0%; Score 40; DB 22; Length 498;				
Best Local Similarity 100.0%; Pred. No. 6.9;				
Matches 7;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1 SSQDYFH 7			
DB	161 ssqdyfh 167			

```

RESULT 7
AA088212
ID AAR88212 standard; Protein; 493 AA.
XX
AC AAR88212;
XX
DT 03-APR-1996 (first entry)
XX
DE Alpha-amylase.
XX
KW Alpha-amylase; thermostable enzyme; baking; Aspergillus niger.
XX
OS Thermomyces lanuginosus CBS 224.63.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= Sig_peptide
XX
PN WO9601323-A1.
XX
PD 18-JAN-1996.
XX
PF 03-JUL-1995; 95WO-EP02607.
XX
PR 04-JUL-1994; 94GB-0013419.
XX
PA (DANI-) DANISCO AS.
XX
PI Michelsen B, Rasmussen P;
XX
DR WPI: 1996-087673/09.
DR N-PSDB; AAT10562.
XX
PT Thermophilic alpha-amylase with activity range of 60-80 degrees C -
PT derived from Thermomyces lanuginosus, useful in the prepn. of
PT foodstuffs and bakery prods. esp. bread
XX
PS Claim 3; Page 36-38; 94pp; English.
XX
CC A thermostable alpha-amylase (AAR88212) of Thermomyces lanuginosus
CC CBS 224.63 is expressed in Aspergillus niger hosts using a gene
CC (AAT10562) isolated from a T. lanuginosus gene library. The
CC recombinant enzyme (54-60 kDa) shows optimal activity at
CC 60-70 deg and pH 5.8-6, has a pi of 3.7 and is active at 60-80
CC deg.
XX
SQ Sequence 493 AA;

```

Query Match 92.5%; Score 37; DB 17; Length 493;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 SSQDYFH 7
Db 161 nsqdyfh 167

```

```

RESULT 8
AA008447
ID AAO08447 standard; Protein; 52 AA.
XX
AC AAO08447;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 22339.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

```

```

XX OS Homo sapiens.
XX
XX PN WO200164835-A2.
XX
XX PD 07-SEP-2001.
XX
XX PF 26-FEB-2001; 2001WO-US04927.
XX
XX PR 28-FEB-2000; 2000US-0515126.
XX
XX PR 18-MAY-2000; 2000US-0577409.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Tang YT, Liu C, Dymnac RT;
XX
XX DR WPI: 2001-514838/56.
XX
XX DR N-PSDB; AAI88378.
XX
XX PT Isolated nucleic acids and polypeptides, useful for preventing
XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
XX PT disorders -
XX
XX PS Claim 20; SEQ ID NO 22339; 1399pp + Sequence listing; English.
XX
XX CC The invention relates to human polynucleotides (AA179941-AA193841) and
XX CC the encoded proteins (AA00010-AA013910) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 52 AA;

```

Query Match 85.0%; Score 34; DB 22; Length 52;
 Best Local Similarity 85.7%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 SSQDYFH 7
Db 2 ssqdyfh 8

```

```

RESULT 9
AAG36453
ID AAG36453 standard; Protein; 1222 AA.
XX
AC AAG36453;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44675.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX OS Arabidopsis thaliana.
XX
XX PN EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-0301439.

```

XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137722.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0143055.
PR 06-JUL-1999; 99US-0143390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 15-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160988.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 1222;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SSQDYFH 7
111111
Db 1130 ssadyfh 1136

RESULT 10
AAG36452
ID AAG36452 standard; Protein; 1257 AA.

XX AAG36452;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 44674.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131448.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0144624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157573.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 1257;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSQDYPH 7
Db 1165 ssadyTh 1171

RESULT 11
AAG36451
ID AAG36451 standard; Protein; 1275 AA.

XX AAG36451;
XX
DT 18-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 44673.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX
PN EP103405-A2.

XX
PD 06-SEP-2000.

XX
PE 25-FEB-2000; 2000EP-0301439.

XX
XX 25-FEB-1999; 99US-0121825.

XX
PR 05-MAR-1999; 99US-0123180.

PR
PR 09-MAR-1999; 99US-0123548.

PR
PR 23-MAR-1999; 99US-0125788.

PR
PR 25-MAR-1999; 99US-0126264.

PR
PR 29-MAR-1999; 99US-0126785.

PR
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0133248.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139765.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140822.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 1275;
Best Local Similarity 85.7%; Pred. No. 2,8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SDDYFH 7
|||
Db 1183 ssadyfh 1189

RESULT 12

AA005772
ID AA005772 standard; Protein: 124 AA.

AC AA005772;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 19664.

Human; cytokine; cell proliferation; cell differentiation; gene therapy;
Human vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
tissue growth factor; immunomodulatory; cancer; leukemia;
nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

DR N-PSDB; AA185703.

Isolated nucleic acids and polypeptides, useful for preventing
diagnosing and treating e.g. leukemia, inflammation and immune
disorders -

Claim 20; SEQ ID NO 19664; 1399bp + Sequence Listing; English.

The invention relates to human polynucleotides (AA179941-AA193841) and
the encoded proteins (AA000010-AA013910) that exhibit activity relating to
cytokine, cell proliferation or cell differentiation or which may induce

production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, hematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activity/inhibit activity and may be useful in the diagnosis and/or
treatment of cancer, leukemia, nervous system disorders, arthritis and
inflammation.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 124 AA:

Query Match 82.5%; Score 33; DB 22; Length 124;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SDDYFH 7
|||
Db 101 sqdyfh 106

RESULT 13

ABG15126
ID ABG15126 standard; Protein: 501 AA.

AC ABG15126;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #15117.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US06631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS79313.

New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -

Claim 20; SEQ ID NO 45485; 103bp; English.

The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
a food supplement. (II) and its binding partners are useful in medical
imaging of sites expressing (II). (I) and (II) are useful for treating

Fri Jun 14 10:36:43 2002

us-09-710-339-2_copy_161_167.rag

Page 12

Search completed: June 13, 2002, 08:53:28
Job time: 278 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:55:51 ; Search time 73.31 Seconds
(without alignments)
9.175 Million cell updates/sec

Title: us-09-710-339-2_COPY_161_167

Perfect score: 40

Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	40	100.0	478	2 JK0201	alpha-amyase (EC
2	40	100.0	498	2 A48305	alpha-amyase (EC
3	40	100.0	499	1 ALAS1	alpha-amyase (EC
4	40	100.0	499	1 ALAS3	alpha-amyase (EC
5	40	100.0	499	2 B48305	alpha-amyase (EC
6	40	100.0	499	2 JS0653	alpha-amyase (EC
7	40	100.0	499	2 JT0466	alpha-amyase (EC
8	40	100.0	499	2 JN0588	alpha-amyase (EC
9	40	100.0	624	1 JC4510	alpha-amyase (EC
10	34	85.0	130	2 H71415	hypothetical prote
11	33	82.5	420	2 T42616	envelope protein -
12	33	82.5	601	2 T18767	hypothetical prote
13	33	82.5	1131	2 T41943	major DNA binding
14	33	82.5	4351	2 T00252	MEGF1 protein - ra
15	32	80.0	144	2 JC7121	hypothetical gland h
16	32	80.0	180	2 B95039	hypothetical prote
17	32	80.0	289	2 A43562	homeotic protein H
18	32	80.0	317	2 E84088	restriction endonu
19	32	80.0	351	2 E86187	YJP8H12.10 (import
20	32	80.0	411	2 S61245	probable vitron g1
21	32	80.0	419	2 T15199	hypothetical prote
22	32	80.0	482	2 A90248	conserved hypochet
23	32	80.0	685	2 S67146	probable membrane
24	32	80.0	820	1 DBECK	thra bifunctional
25	32	80.0	820	2 B90629	aspartokinase I-ho
26	32	80.0	820	2 B85480	aspartokinase I-ho
27	32	80.0	820	2 AC0502	aspartokinase I-ho
28	31	77.5	67	2 T1658	hypothetical prote
29	31	77.5	104	2 S49803	hypothetical prote

30	31	77.5	214	2 B75513	pyridoxamine 5-pho
31	31	77.5	261	2 AF1307	hypothetical prote
32	31	77.5	261	2 AF1679	hypothetical prote
33	31	77.5	386	2 S72435	RNA-binding protei
34	31	77.5	461	2 F86597	UDP-glucose pyroph
35	31	77.5	461	2 H72024	UDP-N-acetylglucos
36	31	77.5	597	2 T35746	hypothetical prote
37	31	77.5	614	1 I48385	hypothetical prote
38	31	77.5	614	2 JC1087	RNA helicase TN22
39	31	77.5	668	2 T10575	RNA helicase, ATP-
40	31	77.5	825	2 S55060	hypothetical prote
41	31	77.5	893	2 T03864	fertilin alpha-II
42	31	77.5	905	2 S55059	hypothetical prote
43	30	75.0	182	2 AE2146	fertilin alpha-I
44	30	75.0	194	2 T22209	hypothetical prote
45	30	75.0	229	2 A89473	hypothetical prote

ALIGNMENTS

```

RESULT 1
JK0201
alpha-amyase (EC 3.2.1.1) - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amyase A
C:Species: Aspergillus oryzae
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JK0201
R: Toda, H.; Kondo, K.; Narita, K.
Proc. Jpn. Acad. 58B, 208-212, 1982
A:Title: The complete amino acid sequence of Taka-amyase A.
A:Reference number: JK0201
A:Accession: JK0201
A:Molecule type: protein
A:Residues: 1-478 <TOD>
C:Comment: One atom of calcium per molecule is essential for the activity.
C:Comment: This enzyme is a glycoprotein.
C:Comment: See also PIR:JT0466 and PIR:JS0240.
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Species: Aspergillus alpha-amyase; alpha-amyase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:173-300/Domain: alpha-amyase core homology <AMY>
F:197/Binding site: carboxylate (Asn) (covalent) #status experimental
F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSQDYFH 7
Db 141 SSQDYFH 147

RESULT 2
A48305
alpha-amyase (EC 3.2.1.1) A precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: A48305
R: Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amyase genes from As
A:Reference number: A48305; MUID:90254827
A:Accession: A48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-498 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

```

A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 40; DB 2; Length 496;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSQDYFH 7
 Db 162 SSQDYFH 168

RESULT 3

ALAS1

alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
 N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2000
 C:Accession: S04548; A33214; J50240; A91930; A93767; A10627
 R:Witzel, S.; Lachmund, A.; Wildhardt, G.; Ruttkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon
 A:Reference number: S04548; MUID:89237897
 A:Accession: S04548
 A:Molecule type: DNA
 A:Residues: 1-499 <WIR>
 A:Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
 A:Genetics: AMY1
 A:Accession: A33214
 A:Molecule type: mRNA
 A:Residues: 1-499 <W12>
 A:Cross-references: GB:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
 R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
 Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
 A:Reference number: J50240; MUID:89378767
 A:Accession: J50240
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A:Genetics: AMY2
 A>Note: the authors refer to this as isozyme II
 R:Isemura, S.; Ikenaka, T.
 J. Biochem. 74, 1-10, 1973
 A:Reference number: A91930; MUID:74001521
 A:Accession: A91930
 A:Molecule type: protein
 A:Residues: 206-225 <ISE>
 R:Narita, K.
 Proc. Jpn. Acad. 51, 285-290, 1975
 A:Reference number: A93767
 A:Accession: A93767
 A:Molecule type: protein
 A:Residues: 434-443,446-447, 'O',449-458, 'GTY',459-464,467-468, 'B',470, 'B',472-499 <NAR>
 R:Matsura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylase A.
 A:Reference number: A37454; MUID:84212370
 A:Contents: annotation: X-ray crystallography, 3.0 angstroms
 R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.
 submitted to the Brookhaven Protein Data Bank, August 1992
 A:Reference number: A51548; PDB:6F7AA
 A:Contents: annotation: X-ray crystallography, 2.1 angstroms, residues 22-497
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: <AMY1>
 A:Gene: amy1
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Genetics: <AMY2>
 A:Gene: amy2
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase 1 #status experimental <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 40; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSQDYFH 7
 Db 162 SSQDYFH 168

RESULT 4

ALAS3

alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
 N:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713
 R:Witzel, S.; Lachmund, A.; Wildhardt, G.; Ruttkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
 A:Reference number: S04549; MUID:89237897
 A:Accession: S04549
 A:Molecule type: DNA
 A:Residues: 1-499 <WIR>
 A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 A:Accession: A33215
 A:Molecule type: translation not shown
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-499 <W12>
 A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
 Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each contain
 A:Reference number: J50240; MUID:89378767
 A:Accession: A44713
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A>Note: the authors refer to this as isozyme I
 R:Matsura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylase A.
 A:Reference number: A37454; MUID:84212370
 A:Contents: annotation: X-ray crystallography, 3.0 angstroms
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: <AMY1>
 A:Gene: amy1
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase 3 #status experimental <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 40; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 |||||||
 Db 162 SSODYFH 168

RESULT 5

alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
 C:Species: Aspergillus awamori
 C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
 C:Accession: B48305
 R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
 Curr. Genet. 17, 203-212, 1990
 A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus
 A:Reference number: A48305; MUID:90254827
 A:Accession: B48305
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499 <KOR>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 C:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 40; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 |||||||
 Db 162 SSODYFH 168

RESULT 6

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
 C:Species: Aspergillus sp.
 C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
 C:Accession: JS0663
 R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
 Biosci. Biotechnol. Biochem. 56, 174-179, 1992
 A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shiroyamii and its expression
 A:Reference number: JS0663; MUID:92323146
 A:Accession: JS0663
 A:Molecule type: mRNA
 A:Residues: 1-499 <SHI>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <AMP>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 40; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 |||||||
 Db 162 SSODYFH 168

Query Match 100.0%; Score 40; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 |||||||
 Db 162 SSODYFH 168

RESULT 7

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 A:Alternate names: glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
 C:Accession: JT0466
 R:Tada, S.; Jimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
 Agric. Biol. Chem. 53, 593-599, 1989
 A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus
 A:Reference number: JT0466
 A:Accession: JT0466
 A:Molecule type: DNA
 A:Residues: 1-499 <TAD>
 C:Comment: See also PIR:JK0201 and PIR:JS0240.
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics:
 A:Inserts: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 40; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 |||||||
 Db 162 SSODYFH 168

RESULT 8

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 A:Alternate names: Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
 C:Accession: JN0588
 R:Tsuakoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsuboi, A.; Ueda, S.
 Gene 84, 319-327, 1989
 A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
 A:Reference number: JN0588; MUID:90128276
 A:Accession: JN0588
 A:Molecule type: mRNA
 A:Residues: 1-499 <TSU>
 C:Comment: The alpha amylases are encoded by multigene family.
 C:Genetics:
 A:Gene: Taa-G1
 A:Inserts: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 40; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7

Db 162 SSODYFH 168

RESULT 9

alpha-amylase (EC 3.2.1.1) precursor - yeast (lipomyces kononenkoae)
N:Alternate names: LKAI protein; raw starch-degrading amylase
N:Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)
C:Species: Lipomyces kononenkoae
C:Date: 07-Feb-1996 #sequence_revision 23-Aug-1996 #text_change 18-Jun-1999
C:Accession: J04510; PC4116
R:Stein, A.J.C.; Marmur, J.; Pretorius, I.S.
Gene 160, 65-71, 1995
A>Title: Cloning, sequence analysis and expression in yeasts of a cDNA containing a lip
A:Reference number: J04510; M01D:96105202
A:Accession: J04510
A:Molecule type: mRNA
A:Residues: 1-624 <STP>
A:Cross-references: GB:U30376; NID:g1173536; PIDN:AA049622.1; PID:g1173537
A:Experimental source: strain IGC4052B
A:Accession: PC4116
A:Molecule type: protein
A:Residues: 29-44 <STP>
A:Experimental source: IGC4052B
C:Genetics:
A:Gene: LKAI
C:Function:
A:Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages
A:Pathway: glycogen/starch degradation
C:Superfamily: lipomyces alpha-amylase; alpha-amylase core homology; glucoamylase starch
C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-624/Product: alpha-amylase #status predicted <MAT>
F:48-141/Domain: glucoamylase starch-binding domain homology <SBD>
F:320-447/Domain: alpha-amylase core homology <AMT>
F:177-185,297-311,387-430,587-622/Disulfide bonds: #status predicted
F:304,344/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:357,377,444/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 40; DB 1; Length 624;
Best Local Similarity 100.0%; Pred. No. 1.8; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSODYFH 7
Db 288 SSODYFH 294

RESULT 10

hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: H71415
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambolt, R.; Weltenegeger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Ehtani, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgomech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reichman, S.; Ans
C.; Chalvatzis, N.
A>Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
A:Reference number: A71400; M01D:98121113
A:Accession: H71415
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-130 <BEV>
A:Cross-references: GB:297338; NID:g2244870; PID:e326912; PID:g2244877
C:Genetics:
A:Map position: 4COP9-4G3845

Query Match 85.0%; Score 34; DB 2; Length 130;
Best Local Similarity 85.7%; Pred. No. 5.6; Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSODYFH 7
Db 38 SSADYFH 44

RESULT 11

envelope protein - equine herpesvirus 4 (strain NS80567)
C:Species: equine herpesvirus 4
A:Variety: strain NS80567
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42616
R:Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A>Title: The DNA sequence of equine herpesvirus-4.
A:Reference number: T22173; M01D:98264497
A:Accession: T42616
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-420 <TEP>
A:Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AA059593.1; PID:g2606021
A:Experimental source: strain NS80567
C:Genetics:
A:Note: 73
C:Superfamily: pseudorabies virus glycoprotein gp63

Query Match 82.5%; Score 33; DB 2; Length 420;
Best Local Similarity 83.3%; Pred. No. 34; Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SODYFH 7
Db 114 NODYFH 119

RESULT 12

hypothetical protein BE10.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18787
R:McMurray, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19022
A:Accession: T18787
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-601 <WIL>
A:Cross-references: EMBL:Z03372; PIDN:CAB07545.1; GSPDB:GN00021; CESP:BE10.3
A:Gene: CESP:BE10.3
A:Map position: 3
A:introns: 21/3; 54/3; 117/2; 157/1; 212/2; 533/3

Query Match 82.5%; Score 33; DB 2; Length 601;
Best Local Similarity 71.4%; Pred. No. 50; Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSODYFH 7
Db 453 NSODYFH 459

RESULT 13

T41943
 major DNA binding protein - human herpesvirus 7 (strain J1)
 C:Species: human herpesvirus 7
 C:Variety: strain J1
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T41943
 R:Nicholas, J.
 submitted to the EMBL Data Library, December 1995
 A:Description: Determination and analysis of the complete nucleotide sequence of human H
 A:Reference number: Z22022
 A:Accession: T41943
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1131 <NIC>
 A:Cross-references: EMBL:U43400; PIDN:AAC54703.1
 A:Experimental source: strain J1
 C:Genetics:
 A:Note: U41
 C:Superfamily: herpesvirus DNA-binding protein

Query Match 82.5%; Score 33; DB 2; Length 1131;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 :|||
 DB 943 ASRDYFH 949

RESULT 14
 T00252
 MEGF1 protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
 C:Accession: T00252
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 Genomics 51, 27-34, 1998
 A>Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
 A:Reference number: Z14126; MUID:98360089
 A:Accession: T00252
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-4351 <NAK>
 A:Cross-references: EMBL:AB011527; NID:93449285; PIDN:BA032458.1; PID:93449286
 A:Experimental source: brain
 C:Genetics:
 A:Gene: MEGF1
 C:Superfamily: rat MEGF1 protein; cadherin repeat homology; EGF homology; laminin G rep
 F:3798-3949/Domain: laminin G repeat homology <LGR>
 F:3953-3985/Domain: EGF homology <EGF>
 F:3992-4023/Domain: EGF homology <EGF1>

Query Match 82.5%; Score 33; DB 2; Length 4351;
 Best Local Similarity 71.4%; Pred. No. 4.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 :|||
 DB 605 NEODYFH 611

RESULT 15
 JC7121
 androgenic gland hormone precursor - common pill bug
 C:Species: Armadillidium vulgare (common pill bug)
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
 C:Accession: JC7121
 R:Okuno, A.; Hasegawa, Y.; Ohira, T.; Katakura, Y.; Nagasawa, H.
 Biochem. Biophys. Res. Commun. 264, 419-423, 1999
 A>Title: Characterization and cDNA cloning of androgenic gland hormone of the terrestria
 A:Reference number: JC7121; MUID:20001935

A:Accession: JC7121
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-144 <OKO>
 A:Cross-references: DDBJ:AB029615; GB:AB029615; NID:96446571; PID:96446572
 A:Experimental source: androgenic gland

Query Match 80.0%; Score 32; DB 2; Length 144;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QDYFH 7
 ||||
 DB 82 QDYFH 86

Search completed: June 13, 2002, 08:55:53
 Job time: 373 sec

Fri Jun 14 10:36:45 2002

us-09-710-339-2_copy_161_167.rpr

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:51 ; search time 38.1 Seconds
(without alignments)
7.114 Million cell updates/sec

Title: US-09-710-339-2_COPY_161_167
Perfect score: 40
Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	40	100.0	498 1	AMVA_ASPAW
2	40	100.0	499 1	AMVA_ASPOR
3	40	100.0	499 1	AMVA_ASPAM
4	40	100.0	499 1	AMVA_ASPSH
5	33	82.5	1131 1	DNBI_HSV7J
6	32	80.0	144 1	AGH_ARMVU
7	32	80.0	289 1	HXDB_MOUSE
8	32	80.0	411 1	VGLM_HSVBC
9	32	80.0	820 1	AKIH_ECOLI
10	32	80.0	1118 1	UBP8_HUMAN
11	31	77.5	104 1	YME8_YEAST
12	31	77.5	198 1	CCG7_HUMAN
13	31	77.5	208 1	LBM_ROME
14	31	77.5	386 1	TIAL_MOUSE
15	31	77.5	614 1	DDX5_HUMAN
16	31	77.5	614 1	DDX5_MOUSE
17	30	75.0	346 1	YDGI_MOUSE
18	30	75.0	432 1	YDGI_ECOLI
19	30	75.0	432 1	YRKO_BACSU
20	30	75.0	623 1	RSDI_YEAST
21	30	75.0	650 1	DDI7_HUMAN
22	30	75.0	714 1	FRDA_HELPJ
23	30	75.0	714 1	FRDA_HELPY
24	30	75.0	937 1	CS32_ECOLI
25	29	72.5	1003 1	SYG_CHLTR
26	29	72.5	144 1	YITE_YEAST
27	29	72.5	217 1	THIE_HELPJ
28	29	72.5	231 1	THIE_HELPY
29	29	72.5	231 1	ALSE_ECOLI
30	29	72.5	242 1	MTGA_KLEPN
31	29	72.5	246 1	PCRB_METH
32	29	72.5	256 1	ERGB_NEUCR
33	29	72.5	312 1	NRTI_CHICK
			374 1	CKR6_HUMAN

34	29	72.5	445 1	YG2V_YEAST
35	29	72.5	675 1	PPK_HELPJ
36	29	72.5	675 1	PPK_HELPY
37	29	72.5	953 1	Y442_HUMAN
38	29	72.5	1101 1	DI42_HUMAN
39	29	72.5	1564 1	PDRA_YEAST
40	29	72.5	1584 1	YJ9G_YEAST
41	28	70.0	106 1	TPIS_STPMU
42	28	70.0	183 1	AAOI_DICDI
43	28	70.0	199 1	TDX_TRYBR
44	28	70.0	200 1	ATKC_AKNSL
45	28	70.0	217 1	HNT2_YEAST

ALIGNMENTS

RESULT 1	AMVA_ASPAW	STANDARD:	PRT:	498 AA.
ID	AMVA_ASPAW			
AC	002905;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase A).			
GN	AMVA.			
OS	Aspergillus awamori.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=105351;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=UVR143F;			
RX	MEDLINE=90254627; PubMed=2340591;			
RA	Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,			
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,			
RA	Berka R.M.;			
RT	*Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus niger var. awamori.;			
RL	Curr. Genet. 17:203-212(1990).			
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides.			
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
DR	EMBL: X52755; CAA36966.1; -			
DR	HSSP: P10529; 7TNA.			
DR	InterPro: IPR00461; Alpha_Amylase.			
DR	Pfam: PF00128; alpha-amylase; 1.			
KW	Glycoprotein; Signal; Multigene family.			
KW	Glycoprotein; Signal; Multigene family.			
FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	498	ALPHA-AMYLASE A.
FT	ACT_SITE	227	227	BY SIMILARITY.
FT	ACT_SITE	231	231	BY SIMILARITY.
FT	ACT_SITE	251	251	BY SIMILARITY.
FT	ACT_SITE	318	318	BY SIMILARITY.
FT	DISULFID	51	59	BY SIMILARITY.
FT	DISULFID	171	185	BY SIMILARITY.
FT	DISULFID	261	304	BY SIMILARITY.
FT	DISULFID	461	496	BY SIMILARITY.
FT	CARBOHYD	218	218	N-LINKED (GLCNAC...) (POTENTIAL).
SQ	SEQUENCE	498 AA:	34880 MW:	763851 IBC01A8A01 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.72;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSODYFH 7
 Db 162 SSODYFH 168

RESULT 2
 AMYA_ASPOR STANDARD; PRT; 499 AA.
 ID AMYA_ASPOR
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 DE (1,4-alpha-D-glucan glucanohydrolase).
 GN AMY1 AND AMY2 AND AMY3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 63503;
 RA Mitsel S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT intron-exon organization.";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89378767; PubMed=2789162;
 RA Genes M.J., Dove M.J., Sellig V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns.";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Toda S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae.";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90128276; PubMed=2612911;
 RA Tsunagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
 RA Uekura S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes.";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A";
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RT A with trypsin and chymotrypsin.";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RT Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; PubMed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution.";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84212370; PubMed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A.";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; PubMed=9283074;
 RA Brzozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution.";
 RL Biochemistry 36:10837-10845(1997).
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I- COPACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I- MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04548; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: J10466; J10466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 7TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR GLYCOSULEDB: P10529; -;
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Hydroxylase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT CONFLICT 93 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTID=CAR_000125.
 O -> R (IN AMY3).
 P -> L (IN AMY3).
 TT -> DC (IN REF. 5).
 Q -> T (IN REF. 5).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 G -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).


```

ID AMY_ASPSH STANDARD; PRT; 499 AA.
AC P30292;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN AMY.
OS Aspergillus shirousami.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eukaryotes; Trichocommataceae; mitosporic Trichocommataceae; Aspergillus.
OX NCBI_TaxID=5070;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92323146; PubMed=1368777;
RA Shibusawa I., Tamura G., Ishikawa T., Hara S.;
RT "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
RT expression in Saccharomyces cerevisiae."
RL Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10461; BAA01255.1; -.
DR PIR: J50663; J50663.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha_amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 499 ALPHA-AMYLASE.
FT ACT_SITE 227 227
FT ACT_SITE 231 231 BY SIMILARITY.
FT ACT_SITE 251 251 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 261 304 BY SIMILARITY.
FT DISULFID 461 496 BY SIMILARITY.
FT CARBOHYD 218 218 N-LINKED (GLCNAC... ) (BY SIMILARITY).
SQ SEQUENCE 499 AA; 54852 MW; 1FB7AE50DA01C03F CRC64;

```

```

Query Match 100.0%; Score 40; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 SSODYFH 7
DB 162 SSODYFH 168

```

```

RESULT 5
DNBI_HSV7J STANDARD; PRT; 1131 AA.
AC P52339;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major DNA-binding protein (MDBP).
GN U41.
OS Human herpesvirus (type 7 / strain J1) (HHV7).

```

```

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-J1;
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
CC REPLICATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U43400; AAC54703.1; -.
DR InterPro: IPR000635; Viral_DNA_bind.
DR Pfam: PF00747; Viral_DNA_bp; 1.
KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
FT ZN-FING 459 475
FT ZN-FING 475 475
SQ SEQUENCE 1131 AA; 129008 MW; 52C97388D1B6D04F CRC64;

```

```

Query Match 82.5%; Score 33; DB 1; Length 1131;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 SSODYFH 7
DB 943 ASRODYFH 949

```

```

RESULT 6
AGH_ARMVU STANDARD; PRT; 144 AA.
AC Q908R2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Androgenic gland hormone precursor.
DE Armadillidium vulgare (woodlice) (Pillbugs).
OS Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Peracarida; Isopoda; Oniscidea; Armadillidiidae;
OC Armadillidium.
OX NCBI_TaxID=13347;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Androgenic gland;
RC MEDLINE=20001935; PubMed=10529379;
RX Okuno A., Hasegawa Y., Ohira T., Katakura Y., Nagasawa H.;
RA "Characterization and cDNA cloning of androgenic gland hormone of the
RA terrestrial isopod Armadillidium vulgare."
RL Biochem. Biophys. Res. Commun. 264:419-423(1999).
RN [2]
RP PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE.
RC TISSUE=Androgenic gland;
RX MEDLINE=99339941; PubMed=10411634;
RA Martin G., Sorokine O., Moniatte M., Bulet P., Hetru C.,
RA Van Dorsselaer A.;
RT "The structure of a glycosylated protein hormone responsible for sex
RT determination in the isopod, Armadillidium vulgare."
RL Eur. J. Biochem. 262:727-736(1999).
CC -1- FUNCTION: CONTROLS SEX DIFFERENTIATION AND THE FORMATION OF MALE
CC APPENDAGES, SPERMATOGENESIS, PIGMENTATION, AND MALE SPECIFIC
CC BEHAVIOR.
CC -1- SUBCELLULAR LOCATION: Secreted.

```



```

CC -1- TISSUE SPECIFICITY: ANDROGENIC GLAND.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: AB029615; BAA6893.1; -.
DR GlycoSiteDB: Q9U8R2; -.
KW Signal; Glycoprotein; Hormone; Sexual differentiation;
KW Spermatogenesis.
FT SIGNAL 1 21
FT CHAIN 22 65 ANDROGENIC GLAND HORMONE B CHAIN.
FT PROPEP 68 113 C PEPTIDE.
FT CHAIN 116 144 ANDROGENIC GLAND HORMONE A CHAIN.
FT DISULFID 33 123 INTERCHAIN (OR 124).
FT DISULFID 42 59 OR 44.
FT DISULFID 44 141 INTERCHAIN (OR 42).
FT DISULFID 124 132 OR 123.
FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (COMPLEX).
FT FTid-CAR.000163.
SQ SEQUENCE 144 AA; 16893 MW; ED1541134628500B CRC64;
QY 3 ODYFH 7
QY 11111
DB 82 ODYFH 86

RESULT 7
HXB8_MOUSE STANDARD; PRT; 289 AA.
ID P23453;
AC 01-NOV-1991 (Rel. 20, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Homeobox protein Hox-D8 (Hox-4.3) (Hox-5.4).
DE HOXD8 OR HOXD-8 OR HOX-4.3.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91209232; PubMed=1982431;
RX Izbisna-Belmonte J.-C., Dolle P., Renucci A., Zappavigna V.,
RX Falkenstein H., Duboule D.;
RA "Primary structure and embryonic expression pattern of the mouse
RA Hox-4.3 homeobox gene.";
RT Development 110:733-745(1990).
RL [2]
RN SEQUENCE OF 191-289 FROM N.A.
RP MEDLINE=91274361; PubMed=1675873;
RX Sadoul R., Featherstone M.;
RA "Sequence analysis of the homeobox-containing exon of the murine
RA Hox-4.3 homeogene.";
RT Biochim. Biophys. Acta 1089:259-261(1991).
RL [3]
RN SEQUENCE OF 195-254 FROM N.A.
RP MEDLINE=92073356; PubMed=1683707;
RX Singh G., Kaur S., Stock J.L., Jenkins N.A., Gilbert D.J.,
RX Copeland N.G., Potter S.S.;
RA "Identification of 10 murine homeobox genes.";
RT Proc. Natl. Acad. Sci. U.S.A. 88:10706-10710(1991).
RL [4]
RN SEQUENCE OF 192-260 FROM N.A.

```

```

RX MEDLINE=92212934; PubMed=1348361;
RA Nazareali A., Kim Y., Nirenberg M.;
RA "Hox-1.11 and Hox-4.9 homeobox genes.";
RT Proc. Natl. Acad. Sci. U.S.A. 89:2883-2887(1992).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEOBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: X56561; CA39911.1; -.
DR EMBL: M87803; AAA37852.1; -.
DR PIR: A38810; A38810.
DR PIR: A41605; A41605.
DR PIR: A43562; A43562.
DR PIR: S16177; S16177.
DR PIR: S15521; S15521.
DR HSSP: P02833; 9ANT.
DR TRANSFAC: T01426; -.
DR MGD: MGI:96209; Hoxd8.
DR InterPro: IPR001827; Antennapedia.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00025; ANTENNAPEDIA.
DR PRINTS: PR00024; HOMEOBOX.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEOX_1.1.
DR PROSITE: PS00032; ANTENNAPEDIA; 1.
DR PROSITE: PS50071; HOMEOBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 15 28 POLY-ALA.
FT DOMAIN 62 89 GLY/PRO-RICH.
FT DOMAIN 108 117 POLY-PRO.
FT DOMAIN 183 188 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 195 254 HOMEOBOX.
FT CONFLICT 207 208 TL -> RV (IN REF. 1).
FT CONFLICT 231 231 T -> S (IN REF. 1).
FT CONFLICT 265 266 EA -> DG (IN REF. 1).
FT CONFLICT 275 275 A -> V (IN REF. 2).
SQ SEQUENCE 289 AA; 31410 MW; 5783099B952BDFE CRC64;
QY 3 ODYFH 7
QY 11111
DB 90 ODYFH 94

RESULT 8
VGLM_HSVBC STANDARD; PRT; 411 AA.
ID VGLM_HSVBC
AC P52370;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Glycoprotein M.
DE GM OR UL10.
OS Bovine herpesvirus type 1 (strain Cooper).
OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OX Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10323;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95313343; PubMed=7793062;
RA Vilek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA Letchworth G.J., Schwyer M.;
RT "Nucleotide sequence analysis of a 30-kb region of the bovine
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT the UL21 to UL4 genes of herpes simplex virus.";
RL Virology 210:100-108(1995).
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN M.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN M.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z48053; CAAB8123.1; -
DR InterPro: IPR000785; Herpes_glycop.
DR Pfam: PF01528; Herpes_glycop. 1.
DR PRINTS: PR00333; HSVINTEGRIMP.
DR Transmembrane; Glycoprotein.
FT TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 89 109 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT CARBOHYD 57 57 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 411 AA; 43629 MW; 20F156DA9FA0158C CRC64;

Query Match 80.0%; Score 32; DB 1; Length 411;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSODYFH 7
DB 295 ASHDFH 301

RESULT 9
AKIH_ECOLI STANDARD: PRT; 820 AA.
ID AC P00561; 047659;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional aspartokinase/homoserine dehydrogenase I (AKI-HDI)
DE [Includes: Aspartokinase I (EC 2.7.2.4); Homoserine dehydrogenase I
DE (EC 1.1.1.3)].
GN THRA OR THRA1 OR THRA2 OR B0002.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
ON [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81077247; PubMed=7003595;
RA Katinka M., Cossart P., Sibilli U., Saint-Girons I., Chalygac M.A.,
RA le Bras G., Cohen G.N., Yaniv M.;
RT "Nucleotide sequence of the thra gene of Escherichia coli.";
RT Proc. Natl. Acad. Sci. U.S.A. 77:5730-5733(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,

```

```

RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. II, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=80077291; PubMed=390305;
RA Cossart P., Katinka M., Yaniv M.;
RT "Construction and expression of a hybrid plasmid containing the
RT Escherichia coli thra and thrB genes.";
RL Mol. Gen. Genet. 175:39-44(1979).
RN [5]
RP SEQUENCE OF 51-129.
RX MEDLINE=80043179; PubMed=387092;
RA Sibilli U., le Bras G., Cossart P., Chalygac M.A., le Bras G.,
RA Briley P.A., Cohen G.N.;
RT "The primary structure of Escherichia coli K 12 aspartokinase I-
RT homoserine dehydrogenase I: sequence of cyanogen bromide peptide CB
RT 3.".
RL Biochimie 61:733-739(1979).
RN [6]
RP REVISION TO 11.
RX MEDLINE=8135751; PubMed=6298218;
RA Cossart P.;
RT Unpublished results, cited by:
RL Zakin M.M., Duchange N., Ferrara P., Cohen G.N.;
RL J. Biol. Chem. 258:3028-3031(1983).
CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P)(+) = L-aspartate 4-
CC semialdehyde + NAD(P)H.
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
CC aspartate.
CC -1- ENZYME REGULATION: THE ENZYME ACTIVITIES ARE REGULATED
CC ALLOSTERICALLY BY L-THREONINE.
CC -1- PATHWAY: CATALYZES 2 NONCONSECUTIVE REACTIONS IN THE COMMON
CC BIOSYNTHETIC PATHWAY LEADING FROM ASP TO DIAMINOPIMELATE AND LYS,
CC TO MET, AND TO THR AND ILE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- MISCELLANEOUS: ASPARTOKINASE II-HOMOSERINE DEHYDROGENASE II AND
CC ASPARTOKINASE III ALSO CATALYZE THE SAME REACTION(S).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE
CC ASPARTOKINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE
CC HOMOSERINE DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V00361; CA233660.1; ALT_SEQ.
DR EMBL: J01706; AAA83914.1; -
DR EMBL: X68872; CAA8734.1; -
DR EMBL: D10483; BAA01286.1; -
DR EMBL: U14003; AAA97301.1; -
DR EMBL: AE000111; AAC73113.1; -
DR EMBL: V00360; CAA23659.1; -
DR EMBL: M28570; AAA24673.1; -
DR EMBL: M10644; AAA24671.1; -
DR PIR: A00671; DEECK.
DR PIR: S40531; S40531.
DR Ecogene; EG10998; thra.

```

```

DR InterPro; IPR002912; ACT.
DR InterPro; IPR001048; Aakkinase.
DR InterPro; IPR001341; Aspartokinase.
DR InterPro; IPR001342; Homoserine_dh.
DR Pfam; PF00696; aakkinase; 1.
DR Pfam; PF01842; ACT; 2.
DR Pfam; PF00742; Homoserine_dh; 1.
DR PROSITE; PS00324; ASPARTOKINASE; 1.
DR PROSITE; PS01042; HOMOSER_DHEXAMINASE; 1.
DR TRANSFERASE; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;
KW Allosteric enzyme; Multifunctional enzyme; Complete proteome.
FT DOMAIN 1 249 ASPARTOKINASE I.
FT DOMAIN 250 470 INTERFERASE.
FT NP_BIND 471 820 HOMOSERINE DEHYDROGENASE I.
FT NP_BIND 471 478 NADP (POTENTIAL).
FT CONFLICT 113 113 Q -> E (IN REF. 5).
FT CONFLICT 230 230 N -> D (IN REF. 3).
FT CONFLICT 375 375 L -> Q (IN REF. 3).
FT CONFLICT 393 393 A -> T (IN REF. 3).
FT CONFLICT 406 406 L -> M (IN REF. 3).
FT CONFLICT 553 553 N -> D (IN REF. 3 AND 4).
FT CONFLICT 587 588 DY -> IT (IN REF. 4).
FT CONFLICT 607 607 I -> T (IN REF. 3).
FT CONFLICT 658 658 R -> T (IN REF. 3).
SQ SEQUENCE 820 AA; 89122 MW; E4DE917A5703F6E7 CRC64;

Query Match
Best local Similarity 80.0%; Score 32; DB 1; Length 820;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSODYFH 7
Db 584 SSMDYH 590

RESULT 10
ID UBP8_HUMAN STANDARD; PRT; 1118 AA.
AC P40818;
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase 8 (EC 3.1.2.15) (Ubiquitin
DE thiolesterase 8) (Ubiquitin-specific processing protease 8)
DE (Deubiquitinating enzyme 8).
GN USP8 OR KIAA0055.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=bone marrow;
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima K.-i., Tabata S.,
RA Seki N., Kawarabayashi Y., Ishikawa K.-i., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1."
RL DNA Res. 1:223-229(1994).
RL -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
ubiquitin + a thiol.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19; ALSO KNOWN AS
FAMILY 2 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; D29956; BAA06225.1; -.
DR MEROPS; C19.011; -.
DR MIM; 603158; -.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR001394; UCH-2.
DR Pfam; PF00581; Rhodanese; 1.
DR Pfam; PF00442; UCH-1; 1.
DR Pfam; PF00443; UCH-2; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS02335; UCH_2_3; 1.
KW Ubiquitin conjugation; Hydrolase; Thiol protease; Multigene family.
FT ACT_SITE 786 786 BY SIMILARITY.
FT ACT_SITE 1059 1059 BY SIMILARITY.
FT ACT_SITE 1067 1067 BY SIMILARITY.
SQ SEQUENCE 1118 AA; 127523 MW; 8B884B7A842P9A9A CRC64;

Query Match
Best local Similarity 80.0%; Score 32; DB 1; Length 1118;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDYFH 7
Db 83 QDYFH 87

RESULT 11
ID YMF8_YEAST STANDARD; PRT; 104 AA.
AC Q04964;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 11.8 kDa protein in OG1-CMA2 intergenic region.
GN YML058W OR YMF958.04.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Devlin K., Churcher C., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; Z46729; CAA86717.1; -.
DR SGD; S0004523; YML058W.
KW Hypothetical protein.
SQ SEQUENCE 104 AA; 11834 MW; 02735C9178E431B2 CRC64;

Query Match
Best local Similarity 77.5%; Score 31; DB 1; Length 104;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
Db 3 SSODYFH 9

RESULT 12
CCG7_HUMAN

```

ID CCG7 HUMAN STANDARD; PRT: 198 AA.
 AC O9BX11:
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Voltage-dependent calcium channel gamma-7 subunit (Neuronal voltage-gated calcium channel gamma-7 subunit).
 DE gated calcium channel gamma-7 subunit).
 GN CACNG7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI-TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21100909; PubMed=11170751;
 RA Burgess D.L., Gefrides L.A., Foreman P.J., Noebels J.L.;
 RT "A cluster of three novel Ca(2+) channel gamma subunit genes on chromosome 19q13.4: evolution and expression profile of the gamma subunit gene family";
 RT Genomics 71:339-350(2001).
 RL Genomics 71:339-350(2001).
 CC -1- FUNCTION: THOUGHT TO STABILIZE THE CALCIUM CHANNEL IN AN INACTIVATED (CLOSED) STATE (BY SIMILARITY).
 CC -1- SUBUNIT: THE L-TYPE CALCIUM CHANNEL IS COMPOSED OF FIVE SUBUNITS: ALPHA-1, ALPHA-2/DELTA, BETA AND GAMMA.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE EMP-22 / EMP / MP20 FAMILY. CACNG SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF288387; AAK20030.1; -
 DR InterPro: IPR004031; PMP22-Claudin.
 DR Pfam: PF00822; PMP22-Claudin; 1.
 KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 KM Calcium channel.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 103 123 POTENTIAL.
 FT TRANSMEM 129 149 POTENTIAL.
 SQ SEQUENCE 198 AA: 22041 MW: AF755430E0A0073B CRC64;
 Query Match 77.5%; Score 31; DB 1; Length 198;
 Best local similarity 71.4%; Pred. No. 18;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SSODYFH 7
 11: 111
 DB 166 SSEOYFH 172
 RESULT 13
 LBM_DROME
 ID LBM_DROME STANDARD; PRT: 208 AA.
 AC Q24188; Q9Y4H0;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Late bloomer protein.
 DE Late bloomer protein.
 GN LBM OR LBL OR CG2374.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI-TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96175586; PubMed=8596956;

RA Kopczyński C.C., Davis G.W., Goodman C.S.;
 RT "A neural tetraspanin, encoded by late bloomer, that facilitates synapse formation.";
 RT Science 271:1867-1870(1996).
 RL Science 271:1867-1870(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.,
 RA Wank R.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miles G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Borkov D., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fester C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Moperson D.,
 RA Merklow G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paciorek J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Pui Y., Reese M.G.,
 RA Ralston R., Remington K., Saunders R.D.C., Scheer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spierias R., Spirding A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassenaar D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: FACILITATES SYNAPSE FORMATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: BELONGS TO THE TETRASPANIN (TM4SF) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U49081; AAA98512.1; -
 DR EMBL: AB003842; AAF59302.1; -
 DR FlyBase: FBgn016032; Lbm.
 DR InterPro: IPR000301; Transmem_4.
 DR Pfam: PF00335; transmembrane4; 1.
 KW Transmembrane.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 41 61 POTENTIAL.
 FT TRANSMEM 67 87 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 SQ SEQUENCE 208 AA: 23005 MW: SEC36525287B7AC9 CRC64;
 Query Match 77.5%; Score 31; DB 1; Length 208;
 Best local similarity 71.4%; Pred. No. 19;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSODYFH 7
1:1111
Db 128 SAODYIH 134

RESULT 14

TITL_MOUSE

ID TITL_MOUSE STANDARD; PRT; 386 AA.

AC P52912;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Nucleolysin TIA-1 (RNA-binding protein TIA-1).

GN TIA1 OR TIA.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RN SEQUENCE FROM N.A.

RP TISSUE=Blood;

RX MEDLINE=96322730; PubMed=8759725;

RA Lowin B., French L., Martinou J.C., Tschoopp J.;

RT "Expression of the CTL-associated protein TIA-1 during murine

embryogenesis.";

RL J. Immunol. 157:1448-1454(1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=97025375; PubMed=8871565;

RA Beck A.R.P., Medley O.G., O'Brien S., Anderson P., Streuli M.;

RT "Structure, tissue distribution and genomic organization of the

murine RRM-type RNA binding proteins TIA-1 and TIAR.";

RL Nucleic Acids Res. 24:3829-3836(1996).

CC -!- FUNCTION: RNA-BINDING PROTEIN. POSSESSES NUCLEOLYTIC ACTIVITY

AGAINST CYTOTOXIC LYMPHOCYTE TARGET CELLS. MAY BE INVOLVED IN

APOPTOSIS (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES OF CYTOLYTIC

T-LYMPHOCYTES (BY SIMILARITY).

CC -!- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).

CC -!- SIMILARITY: TO NUCLEAR RNA-BINDING PROTEINS. STRONGEST TO TIAR.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U00689; AAA03711.1; -

DR EMBL: U55867; AAC52871.1; -

DR MGD: MGI:107914; Tia1.

DR InterPro: IPR000504; RRM.

DR Pfam: PF00076; rrm. 3.

DR SMART: SM00360; RRM; 3.

DR PROSITE: PS0102; RRM; 3.

DR PROSITE: PS00030; RRM_RNP_1; 2.

KW RNA-binding; Apoptosis; Repeat.

FT DOMAIN 7 83 RNA-BINDING (RRM) 1.

FT DOMAIN 106 184 RNA-BINDING (RRM) 2.

FT DOMAIN 214 286 RNA-BINDING (RRM) 3.

SO SEQUENCE 386 AA; 42800 MW; 51BFEE90E046D3AE CRC64;

Query Match 77.5%; Score 31; DB 1; Length 386;
Best Local Similarity 83.3%; Pred. No. 37;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSODYFH 7
1111111
Db 102 SSODYFH 107

RESULT 15

IDDX5_HUMAN

STANDARD; PRT; 614 AA.

AC P17844;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable RNA-dependent helicase p68 (DEAD-box protein p68) (DEAD-box

protein 5).

GN DDX5 OR HLR1 OR HELR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxId=9606;

RN SEQUENCE FROM N.A.

RX MEDLINE=90272412; PubMed=2349099;

RA Hloch P., Stahl H.;

RT "Complete cDNA sequence of the human p68 protein.";

RL Nucleic Acids Res. 18:3045-3045(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=91141480; PubMed=1996094;

RA 1990 R.D., Jamieson D.D., McNeill S.A., Southgate J., McPheat J.,

Lane D.P.;

RT "p68 RNA helicase: identification of a nucleolar form and cloning of

related genes containing a conserved intron in yeasts.";

RL Mol. Cell. Biol. 11:1326-1333(1991).

RN [3]

RP SEQUENCE FROM N.A.

RA Hloch P., Roessler O.G., Weitzenecker T., Stahl H.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE OF 21-614 FROM N.A.

RX MEDLINE=88189326; PubMed=2451786;

RA Ford M.J., Anton I.A., Lane D.P.;

RT "Nuclear protein with sequence homology to translation initiation

factor eIF-4A.";

RL Nature 332:736-738(1988).

CC -!- FUNCTION: RNA-DEPENDENT ATPASE ACTIVITY. THE RATE OF ATP

HYDROLYSIS IS HIGHLY STIMULATED BY SINGLE-STRANDED RNA.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX5/DDX17

SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X52104; CAA6324.1; -

DR EMBL: AF015812; AAB84094.1; -

DR EMBL: X15729; CAA33751.1; -

DR PIR: S06377; S06377.

DR PIR: S10181; S10181.

DR PIR: S14045; S14045.

DR Aarhus/Chent-ZPAGE; 1605; NEPHGE.

DR HSP; Q58083; IHV8.

DR MIM; 180630; -

DR InterPro: IPR001410; DEAD.

DR InterPro: IPR000629; DEAD_ATP_helicase.

DR InterPro: IPR001650; Helicase_C.

DR Pfam: PF00270; DEAD; 1.

DR Pfam: pf00271; helicase_C; 1.

DR SMART: SM00487; DEXDC; 1.

DR SMART: SM00490; HELIC; 1.

DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.

DR ATP-binding; RNA-binding; Helicase; Nuclear protein.

NP BIND 138 145 ATP (BY SIMILARITY).
 SITE 248 251 DEAD BOX
 SQ SEQUENCE 614 AA; 69148 MW; 84DF684FD6871594 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 614;
 Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSQDYEH 7
 11:111
 Db 421 SSQDYEH 427

Search completed: June 13, 2002, 09:11:53
 Job time: 1033 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:10 ; Search time 125.81 Seconds
(without alignments)
9.625 Million cell updates/sec

Title: US-09-710-339-2_COPY_161_167

Perfect score: 40
Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriophage:*
17: sp.archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	40	100.0	499	3	Q96TH4	Q96TH4 aspergillus
2	40	100.0	624	3	Q01117	Q01117 lipomyces k
3	35	87.5	760	5	Q9XJ00	Q9XJ00 giardia lam
4	34	85.0	130	10	Q23373	Q23373 arabidopsis
5	33	82.5	420	12	O55525	O55525 equine herp
6	33	82.5	420	12	O39309	O39309 equine herp
7	33	82.5	585	11	Q9D718	Q9D718 mus musculu
8	33	82.5	601	5	O17552	O17552 caenorhabdi
9	33	82.5	1131	12	O56282	O56282 human herpe
10	33	82.5	1731	4	O9P230	O9P230 homo sapien
11	33	82.5	3124	4	O96L91	O96L91 homo sapien
12	33	82.5	4351	11	O98277	O98277 rattus norv
13	32	80.0	163	11	O9C0Q8	O9C0Q8 mus musculu
14	32	80.0	180	16	Q97SK2	Q97SK2 streptococc
15	32	80.0	215	14	P97975	P97975 unidentified
16	32	80.0	218	11	Q9C0H8	Q9C0H8 mus musculu

17	32	80.0	317	16	Q9K764	Q9K764 bacillus ha
18	32	80.0	351	10	Q23041	Q23041 arabidopsis
19	32	80.0	419	5	O01765	O01765 caenorhabdi
20	32	80.0	438	12	O39493	O39493 bovine herp
21	32	80.0	482	17	O97ZF2	O97ZF2 sulfolobus
22	32	80.0	529	5	Q9YK03	Q9YK03 drosophila
23	32	80.0	575	11	Q91ZG4	Q91ZG4 mus musculu
24	32	80.0	583	4	Q92802	Q92802 homo sapien
25	32	80.0	685	3	Q08683	Q08683 saccharomyc
26	31	77.5	67	12	Q84487	Q84487 paramecium
27	31	77.5	172	11	Q90XU1	Q90XU1 mesocricetu
28	31	77.5	172	11	Q99MG7	Q99MG7 mesocricetu
29	31	77.5	200	6	Q28465	Q28465 macropus eu
30	31	77.5	214	4	Q96B58	Q96B58 homo sapien
31	31	77.5	214	16	Q9RX20	Q9RX20 delnoccocus
32	31	77.5	219	6	Q95LS2	Q95LS2 macaca fasc
33	31	77.5	261	16	Q92AE6	Q92AE6 listeria in
34	31	77.5	304	11	Q9CSD9	Q9CSD9 mus musculu
35	31	77.5	388	13	Q90WNO	Q90WNO xenopus lae
36	31	77.5	461	16	Q92750	Q92750 chlamydia p
37	31	77.5	480	3	Q90Y07	Q90Y07 emericella
38	31	77.5	595	13	Q9W744	Q9W744 gallus gall
39	31	77.5	597	2	Q952K2	Q952K2 streptomyce
40	31	77.5	608	13	Q918S3	Q918S3 xenopus lae
41	31	77.5	668	10	Q91DV7	Q91DV7 arabidopsis
42	31	77.5	761	5	Q9NLA0	Q9NLA0 asterina pe
43	31	77.5	825	6	Q28477	Q28477 macaca fasc
44	31	77.5	838	6	O19056	O19056 papio anudi
45	31	77.5	893	5	O16221	O16221 caenorhabdi

ALIGNMENTS

RESULT 1
ID Q96TH4 PRELIMINARY; PRT; 499 AA.
AC Q96TH4;
DF 01-DEC-2001 (TREMREL. 19, Created)
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
GN AMYA OR AMY1.
GN Aspergillus oryzae, and
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
CX NCBI_TaxID=5062, 5059;
CX [1]
RN SEQUENCE FROM N.A.
RP SPECIES=A.oryzae; STRAIN=R1840;
RC MEDLINE=2028310; PubMed=10830498;
RA Gomi K., Akeno T., Minekoki T., Ozeki K., Kumagai C., Okazaki N.,
RA Iimura Y.,
RT "Molecular cloning and characterization of a transcriptional activator
RT gene, amyR, involved in the amyloid gene expression in Aspergillus
RT oryzae.";
RL Biosci. Biotechnol. Biochem. 64:816-827(2000).
RN [2]
RN SEQUENCE FROM N.A.
RP SPECIES=A.flavus; STRAIN=86-10D;
RC Fakhoury A.M., Woloshuk C.P.,
RA "Amy1, the alpha-amylase gene of Aspergillus flavus: Involvement in
RT aflatoxin biosynthesis in maize kernels.";
RT Phytopathology 89:908-914(1999).
RL EMBL; AB021876; BAA95703.1; -;
DR EMBL; AF139925; AAF14264.1; -;
SQ SEQUENCE 499 AA; 54804 MW; EEFA2ADA71D20DA9 CRC64;

Query Match 100.0%; Score 40; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
 |||||
 Db 162 SSODYFH 168

RESULT 2
 ID 001117 PRELIMINARY; PRT; 624 AA.
 AC 001117;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1)
 DE (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE) (LKA1).
 GN LKA1.
 OS Lipomyces kononenkoae.
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Lipomycetaceae; Lipomyces.
 OK NCBI_TaxID=34357;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=IGC4052B;
 RX MEDLINE=96105202; PubMed=8529895;
 RA Steyn A.J.C., Marmur J., Pretorius I.S.;
 RT "Cloning, sequence analysis and expression in yeasts of a cDNA
 RL containing a Lipomyces kononenkoae alpha-amylase-encoding gene."
 RL Gene 166:65-71(1995).
 RN [2]
 RP SEQUENCE OF 29-44.
 RC STRAIN=IGC4052B;
 RX MEDLINE=96133108; PubMed=8593683;
 RA Steyn A.J.C., Pretorius I.S.;
 RT "Characterization of a novel alpha-amylase from Lipomyces kononenkoae
 RT and expression of its gene (LKA1) in Saccharomyces cerevisiae."
 RL Curr. Genet. 28:526-533(1995).
 CC -1- FUNCTION: LIBERATES REDUCING GROUPS FROM POLYMERS CONTAINING BOTH
 CC 1,4-ALPHA AND 1,6-ALPHA BONDS.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL: U30376; AAC49622.1; ALT_INIT.
 DR HSP; P10529; 77NA.
 DR InterPro: IPR00461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Calcium;
 KW Glycoprotein.
 FT SIGNAL 1 28
 FT CHAIN 29 624 ALPHA-AMYLASE.
 FT ACT_SITE 353 353 BY SIMILARITY.
 FT ACT_SITE 357 357 BY SIMILARITY.
 FT ACT_SITE 377 377 BY SIMILARITY.
 FT ACT_SITE 444 444 BY SIMILARITY.
 FT DISULFID 177 185 BY SIMILARITY.
 FT DISULFID 297 311 BY SIMILARITY.
 FT DISULFID 387 430 BY SIMILARITY.
 FT DISULFID 587 622 BY SIMILARITY.
 FT CARBOHYD 304 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 624 AA; 68876 MW; 87EB16534F5A9A9F CRC64;

Query Match 100.0%; Score 40; DB 3; Length 624;
 Best Local Similarity 100.0%; Pred. No. 3.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
 |||||
 Db 288 SSODYFH 294

RESULT 3
 ID 09XZJ0 PRELIMINARY; PRT; 760 AA.
 AC 09XZJ0;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ANKYRIN-LIKE PROTEIN.
 DE Giardia lamblia (Giardia intestinalis).
 OS Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OK NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PORTLAND;
 RA Sekyere E.O., Edwards M.R., Stewart T.S., Schofield P.J., Knodler L.;
 RT "Ankyrin-like protein from Giardia intestinalis."
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF144322; AAD28486.1; -.
 DR InterPro: IPR002110; ANK.
 DR Pfam: PF00023; ank; 2.
 DR SMART: SM00248; ANK; 1.
 KW ANK repeat; Repeat.
 SQ SEQUENCE 760 AA; 85542 MW; 12B54841019CD475 CRC64;

Query Match 87.5%; Score 35; DB 5; Length 760;
 Best Local Similarity 85.7%; Pred. No. 45;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
 |||||
 Db 5 SSODYFH 11

RESULT 4
 ID 023373 PRELIMINARY; PRT; 130 AA.
 AC 023373;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE HYPOTHETICAL PROTEIN.
 DE AT4G15190.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terry N.,
 RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
 RA Pulgomech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
 RA Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,
 RA Schueller C., Chalvatiz N.,
 RL submitted (Jul-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z97338; CAB10298.1; -.
 DR EMBL: AL161540; CAB78561.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 130 AA; 14962 MW; DD87D2003030184 CRC64;

Query Match 85.0%; Score 34; DB 10; Length 130;
 Best Local Similarity 85.7%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSODYFH 7

DB 38 SSADYFH 44

RESULT 5
ID 055525 PRELIMINARY; PRT; 420 AA.

AC 055525;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE GLYCOPROTEIN I HOMOLOGUE.
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TH20;
RA MEDLINE=98185635; PubMed=9524947;
RA Damiani A.M., Matsumura T., Yokoyama N., Maeda K., Miyazawa T.,
Kai C., Mikami T.;
RT "Nucleotide sequences of glycoprotein I and E genes of equine
herpesvirus type 4.";
RL J. Vet. Med. Sci. 60:219-225(1998).
DR EMBL: AB005749; BAA25021.1;
DR InterPro: IPR002874; Herpes-G1.
DR Pfam: PF01688; Herpes-G1; 1.
SQ SEQUENCE 420 AA; 45696 MW; 1F64F5E179C2D916 CRC64;

Query Match 82.5%; Score 33; DB 12; Length 420;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SDDYFH 7
DB 114 NDDYFH 119

RESULT 6
ID 039309 PRELIMINARY; PRT; 420 AA.
AC 039309;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE COUNTERPART OF HSV-1 GENE US7 AND VZV GENE 67.
GN 73.
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NS80567;
RC MEDLINE=98264497; PubMed=9603335;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-4.";
RL J. Gen. Virol. 79:1197-1203(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NS80567;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF030027; AAC59593.1;
DR InterPro: IPR002874; Herpes-G1.
DR Pfam: PF01688; Herpes-G1; 1.
SQ SEQUENCE 420 AA; 45710 MW; 1EB944825DF3D852 CRC64;

Query Match 82.5%; Score 33; DB 12; Length 420;
Best Local Similarity 83.3%; Pred. No. 62;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SDDYFH 7
DB 114 NDDYFH 119

RESULT 7
ID 09D718 PRELIMINARY; PRT; 585 AA.
AC 09D718;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 2310007D09R1K PROTEIN.
GN 2310007D09R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
Radota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schröml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Williams L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK009199; BAB26134.1;
DR MGD: MGI:1919128; 2310007D09R1K.
DR InterPro: IPR001412; tRNA-synt.1.
DR PROSITE: PS00178; AA.tRNA.LIGASE.I; UNKNOWN.1.
SQ SEQUENCE 585 AA; 64324 MW; E3528EDAAE13FBFF CRC64;

Query Match 82.5%; Score 33; DB 11; Length 585;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDDYFH 7
DB 366 SDDYFH 372

RESULT 8
ID 017552 PRELIMINARY; PRT; 601 AA.
AC 017552;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE BE10.3 PROTEIN.
GN BE10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;

RN [1]
 RP SEQUENCE FROM N.A.
 RL McMurtry A.A.; 1997) to the EMBL/GenBank/DBJ databases.
 RN Submitted (Mar-1997) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RX MEDLINE=99069613; PubMed=9851916;
 RP SEQUENCE FROM N.A.
 RX none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology";
 RL Science 283:2012-2018(1998).
 DR EMBL; 293372; CAB07545.1; -;
 SQ SEQUENCE 601 AA; 69490 MW; 30084EEBEC98B7E9 CRC64;

Query Match 82.5%; Score 33; DB 5; Length 601;
 Best Local Similarity 71.4%; Pred. No. 92;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
 :|:|:|:|
 DB 453 NSRDYFH 459

RESULT 9
 ID 056282 PRELIMINARY; PRT; 1131 AA.
 AC 056282;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SINGLE-STRANDED DNA-BINDING PROTEIN.
 GN U41.
 OS Human herpesvirus 7.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae.
 OX NCBI_TaxID=10372;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RK;
 RX MEDLINE=9840941; PubMed=9581785;
 RA Megaw A.G., Rapaport D., Avidor B., Frenkel N., Davison A.J.;
 RT "The DNA sequence of the RK strain of human herpesvirus 7.";
 RL Virology 244:119-132(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RK;
 RA Megaw A.G., Frenkel N., Davison A.J.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF037218; AAC40755.1; -;
 DR Interpro; IPR000635; Viral_DNA_bind.
 DR Pfam; PF00747; Viral_DNA_dp; 1.
 KW DNA-binding.
 SQ SEQUENCE 1131 AA; 129018 MW; 165B3AAC703CC82D CRC64;

Query Match 82.5%; Score 33; DB 12; Length 1131;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
 :|:|:|:|
 DB 943 ASRDYFH 949

RESULT 10
 ID 09P230 PRELIMINARY; PRT; 1731 AA.
 AC 09P230;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE KIAA1498 PROTEIN (FRAGMENT).

GN KIAA1498.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20277482; PubMed=10819331;
 RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human
 RT genes. XVII. The complete sequences of 100 new cDNA clones from brain
 RT which code for large proteins in vitro";
 RL DNA Res. 7:143-150(2000).
 DR EMBL; AB040931; BAA96022.1; -;
 DR Interpro; IPR000330; SNEF2_N.
 DR Pfam; PF00176; SNEF2_N; 2.
 DR NON_TER 1
 FT NON_TER 1731 1731
 SQ SEQUENCE 1731 AA; 188228 MW; D628E95F1021A8F CRC64;

Query Match 82.5%; Score 33; DB 4; Length 1731;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SODYFH 7
 :|:|:|:|
 DB 1250 SODYFH 1255

RESULT 11
 ID 096L91 PRELIMINARY; PRT; 3124 AA.
 AC 096L91;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE P400 SWI2/SNF2-RELATED PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21400441; PubMed=11509179;
 RA Fuchs M., Gerber J., Drapkin R., Sif S., Ikura T., Ogrzyzko V.,
 RA Lane W.S., Nakatani Y., Livingston D.M.;
 RT "The p400 complex is an essential E1A transformation target";
 RL Cell 106:297-307(2001).
 DR EMBL; AY044869; AAK97789.1; -;
 SQ SEQUENCE 3124 AA; 340146 MW; E8F57D6C7BD01B9 CRC64;

Query Match 82.5%; Score 33; DB 4; Length 3124;
 Best Local Similarity 83.3%; Pred. No. 5.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SODYFH 7
 :|:|:|:|
 DB 1249 SODYFH 1254

RESULT 12
 ID 088277 PRELIMINARY; PRT; 4351 AA.
 AC 088277;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MEGF1.
 GN MEGF1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX MEDLINE=98360089; PubMed=9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RL like motifs by motif-trap screening.";
RT Genomics 51:27-34(1998).
CC -1 SIMILARITY: CONTAINS 33 CADHERIN DOMAINS.
DR EMBL: AB011527; BAA32458.1; -.
DR HSSP: P08703; IBE9.
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001791; laminin_G.
DR Pfam: PF000028; cadherin; 33.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00054; laminin_G; 1.
DR PRINTS: PR00205; CADHERIN.
DR SMART: SM00112; CA; 33.
DR SMART: SM00181; EGF; 2.
DR SMART: SM00282; LamG; 1.
DR PROSITE: PS00232; CADHERIN_1; 1.
DR PROSITE: PS00268; CADHERIN_2; 33.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 2.
DR Calcium-binding; Cell adhesion; BGF-like domain; Glycoprotein.
SQ SEQUENCE 4351 AA; 480644 MW; C564E7F98BB388 CRC64;

Query Match      82.5%; Score 33; DB 11; Length 4351;
Best Local Similarity 71.4%; Pred. No. 7.9e+00;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SQDYFH 7
   : |||||
Db 605 NEODYFH 611

RESULT 13
ID 09CU08 PRELIMINARY; PRT; 163 AA.
AC 09CU08;
DT 01-JUN-2001 (TRENBLREL. 17, Created)
DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
DE 4921540P06RIK (TRENBLREL. 19, Last annotation update)
DE 4921540P06RIK PROTEIN (FRAGMENT).
GN 4921540P06RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kiehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsi G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

```

```

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK015019; BAB2676.1; -.
DR HSSP: P14653; IBE72.
DR MGD: MGI:1913985; 4921540P06RIK.
DR InterPro: IPR001827; Antennapedia.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00025; ANTENNAPEDIA.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00032; ANTENNAPEDIA; UNKNOWN_1.
DR PROSITE: PS00071; HOMEBOX_2; 1.
FT NON_TER 1
SQ SEQUENCE 163 AA; 18199 MW; E956CF5E76D1E8D1 CRC64;

```

```

Query Match      80.0%; Score 32; DB 11; Length 163;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 QDYFH 7
   : |||||
Db 11 QDYFH 15

```

```

RESULT 14
ID 097SK2 PRELIMINARY; PRT; 180 AA.
AC 097SK2;
DT 01-OCT-2001 (TRENBLREL. 18, Created)
DT 01-OCT-2001 (TRENBLREL. 18, Last sequence update)
DE 4921540P06RIK (TRENBLREL. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN SP0332.
GN SP0332.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Urdarack T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Anguoli S., Dickinson T., Hickey E.K.,
RA Holt J.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL: AE007345; AAK74507.1; -.
DR TIGR: SP0332; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 180 AA; 20652 MW; E5024C611428B844 CRC64;

```

```

Query Match      80.0%; Score 32; DB 16; Length 180;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 QDYFH 7
   : |||||
Db 136 QDYFH 140

```

```

RESULT 15
P97975

```

ID P97975 PRELIMINARY; PRT: 215 AA.
 AC P97975;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 70-KDA HEAT SHOCK PROTEIN (FRAGMENT).
 GN HSP70.
 OS unidentified soil organism.
 OC unclassified: environmental samples.
 OX NCBI_TaxID=46465;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yap W.H., Li X., Soong T.W., Davies J.E.;
 RT "Genetic diversity of soil microorganisms assessed by analysis of
 RT hsp70 (dnaK) sequences.";
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U49144; AAB48232.1; -;
 DR InterPro: IPR01023; HSP70.
 DR Pfam: PF00012; HSP70.1
 DR PRINTS: PR00301; HEATSHOCK70.
 DR PROSITE: PS00329; HSE70_2; 1.
 DR PROSITE: PS01036; HSE70_3; 1.
 KW Heat shock.
 FT NON_TER 1 1
 FT NON_TER 215 215
 SQ SEQUENCE 215 AA; 23694 MW; 9146A64A3AC9623D CRC64;

Query Match 80.0%; Score 32; DB 14; Length 215;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDYFH 7
 11111
 Db 196 QDYFH 200

Search completed: June 13, 2002, 09:11:12
 Job time: 1057 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:30 ; Search time 57.95 seconds
(without alignments)
2.950 Million cell updates/sec

Title: US-09-710-339-2_COPY_161_167
Perfect score: 40
Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	40	100.0	478 1 US-08-720-899-7	Sequence 7, Appli
2	40	100.0	478 1 US-08-459-610-7	Sequence 7, Appli
3	40	100.0	478 2 US-08-343-804-7	Sequence 7, Appli
4	40	100.0	478 2 US-08-339-715A-2	Sequence 2, Appli
5	40	100.0	478 2 US-08-600-908A-10	Sequence 10, Appli
6	40	100.0	478 3 US-08-683-838A-10	Sequence 10, Appli
7	40	100.0	478 3 US-09-182-859-7	Sequence 7, Appli
8	32	80.0	369 1 US-08-700-359-21	Sequence 21, Appli
9	32	80.0	583 2 US-08-616-392C-4	Sequence 4, Appli
10	32	80.0	820 2 US-08-380-182-23	Sequence 23, Appli
11	30	75.0	25 1 US-08-485-455D-41	Sequence 41, Appli
12	30	75.0	25 2 US-08-482-130C-41	Sequence 41, Appli
13	30	75.0	25 2 US-08-484-211C-41	Sequence 41, Appli
14	30	75.0	25 3 US-08-906-769-41	Sequence 41, Appli
15	30	75.0	25 3 US-08-906-616-41	Sequence 41, Appli
16	30	75.0	25 4 US-08-817-795-41	Sequence 41, Appli
17	30	75.0	25 4 US-08-485-443B-41	Sequence 41, Appli
18	30	75.0	25 4 US-08-639-075A-11	Sequence 41, Appli
19	30	75.0	25 4 US-09-012-431-41	Sequence 41, Appli
20	30	75.0	25 4 US-09-012-692-41	Sequence 41, Appli
21	30	75.0	25 4 US-08-906-613-41	Sequence 41, Appli
22	30	75.0	25 5 PCT-US95-14442A-41	Sequence 41, Appli
23	30	75.0	64 1 US-08-485-455D-75	Sequence 75, Appli
24	30	75.0	64 2 US-08-482-130C-75	Sequence 75, Appli
25	30	75.0	64 2 US-08-484-211C-75	Sequence 75, Appli
26	30	75.0	64 3 US-08-906-769-75	Sequence 75, Appli
27	30	75.0	64 3 US-08-906-616-75	Sequence 75, Appli

28	30	75.0	64 4 US-08-817-795-75	Sequence 75, Appli
29	30	75.0	64 4 US-08-485-443B-75	Sequence 75, Appli
30	30	75.0	64 4 US-08-639-075A-75	Sequence 75, Appli
31	30	75.0	64 4 US-09-012-431-75	Sequence 75, Appli
32	30	75.0	64 4 US-09-012-692-75	Sequence 75, Appli
33	30	75.0	64 4 US-08-906-613-75	Sequence 75, Appli
34	30	75.0	64 5 PCT-US95-14442A-75	Sequence 75, Appli
35	30	75.0	223 1 US-08-485-455D-17	Sequence 17, Appli
36	30	75.0	223 2 US-08-482-130C-17	Sequence 17, Appli
37	30	75.0	223 2 US-08-484-211C-17	Sequence 17, Appli
38	30	75.0	223 4 US-08-817-795-17	Sequence 17, Appli
39	30	75.0	223 4 US-08-485-443B-17	Sequence 17, Appli
40	30	75.0	223 5 PCT-US95-14442A-17	Sequence 17, Appli
41	30	75.0	224 3 US-08-906-769-17	Sequence 17, Appli
42	30	75.0	224 3 US-08-906-616-17	Sequence 17, Appli
43	30	75.0	224 4 US-08-639-075A-17	Sequence 17, Appli
44	30	75.0	224 4 US-09-012-431-17	Sequence 17, Appli
45	30	75.0	224 4 US-09-012-692-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowrey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054,214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7
Query Match 100.0%; Score 40; DB 1; Length 478;
Best local Similarity 100.0%; Pred. No 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
111111
DB 141 SSODYFH 147

RESULT 2
US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; MEDIUM TYPE: Floppy disk
; COMPUTER READABLE FORM:
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459, 610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-459-610-7

Query Match 100.0%; Score 40; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
111111
DB 141 SSODYFH 147

RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5850837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: No. 58308370 No. 5830837disk of No. 5830837th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,804
FILING DATE: 22-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
; US-08-343-804-7

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
111111
DB 141 SSODYFH 147

RESULT 4
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michiyo
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; TITLE OF INVENTION: NEOPULULINASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: POPHAM, HARK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; STREET: G. Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; MEDIUM TYPE: Floppy disk
; COMPUTER READABLE FORM:
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Melser, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
Db 141 SSODYFH 147

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-600-908A-10

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
Db 141 SSODYFH 147

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-683-838A-10

Query Match 100.0%; Score 40; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
Db 141 SSODYFH 147

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgaard-Frantzen, Henrik

;; TITLE OF INVENTION: Alpha-Amylase Mutants
;; FILE REFERENCE: 4796.204-US
;; CURRENT APPLICATION NUMBER: US/09/182,859
;; EARLIER FILING DATE: 1998-10-29
;; EARLIER APPLICATION NUMBER: 0515/96
;; EARLIER FILING DATE: 1996-04-30
;; EARLIER APPLICATION NUMBER: 0712/96
;; EARLIER FILING DATE: 1996-06-28
;; EARLIER APPLICATION NUMBER: 0775/96
;; EARLIER FILING DATE: 1996-07-11
;; EARLIER APPLICATION NUMBER: 1263/96
;; EARLIER FILING DATE: 1996-11-08
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 7
;; LENGTH: 478
;; TYPE: PRT
;; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 40; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
DB 141 SSODYFH 147

RESULT 8
US-08-700-359-21
; Sequence 21, Application US/08700359
; Patent No. 5766925
; GENERAL INFORMATION:
; APPLICANT: SUGIMOTO, MASAKAZU
; APPLICANT: USUDA, YOSHIHIRO
; APPLICANT: SUZUKI, TOMOKO
; APPLICANT: TANAKA, AKIKO
; APPLICANT: MATSUI, HIROSHI
; TITLE OF INVENTION: METHOD OF PRODUCING L-LYSINE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OHLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS-DOS TEXT EDITOR
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,359
; FILING DATE: 08-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP-35019
; FILING DATE: 04-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: NORMAN F. OHLON
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-819-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2230
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-08-700-359-21

Query Match 80.0%; Score 32; DB 1; Length 369;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSODYFH 7
DB 133 SSMDYFH 139

RESULT 9
US-08-616-392C-4
; Sequence 4, Application US/08616392C
; Patent No. 5998165
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerbloom, Ingrid E.
; APPLICANT: Sellhammer, Jeffrey
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCLIA
; TITLE OF INVENTION: AND PANCLIA ASSOCIATED WITH PANCREATIC CANCER
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,392C
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/581,240
; FILING DATE: 29-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0052-105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: hnt
; CLONE: 496071
US-08-616-392C-4

Query Match 80.0%; Score 32; DB 2; Length 583;
Best Local Similarity 71.4%; Pred. No. 87;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSODYFH 7
DB 433 STDDYFH 439

RESULT 10
US-08-380-182-23
; Sequence 23, Application US/08380182


```
; Patent No. 5858749
; GENERAL INFORMATION:
; APPLICANT: Matthews, Benjamin F.
; APPLICANT: Weisemann, Jane M.
; TITLE OF INVENTION: A Bifunctional Protein From Carrots
; TITLE OF INVENTION: (Daucus carota) with Aspartokinase and Homoserine
; TITLE OF INVENTION: Dehydrogenase Activities"
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janelle S. Graeter
; STREET: Room 411, Bldg. 005, BARC-W
; CITY: Beltsville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20705
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/380,182
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Graeter, Janelle S.
; REGISTRATION NUMBER: 35,024
; REFERENCE/DOCKET NUMBER: 0226.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-504-6629
; TELEFAX: 301-504-5060
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 820 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
;
US-08-380-182-23
```

```
Query Match 80.0%; Score 32; DB 2; Length 820;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 SSODYEH 7
Db 584 SSMDYIH 590

RESULT 11
US-08-485-455D-41
; Sequence 41, Application US/08485455D
; Patent No. 5712143
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary L.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,455D
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C1-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-485-455D-41
```

```
Query Match 75.0%; Score 30; DB 1; Length 25;
Best Local Similarity 66.7%; Pred. No. 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 SODYFH 7
Db 20 AEDYFH 25
```

```
RESULT 12
US-08-482-130C-41
; Sequence 41, Application US/08482130C
; Patent No. 5962257
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary L.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,130C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-130C-41

Query Match
Best Local Similarity 75.0%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 SODYFH 7
:|||||
DB 20 AEDYFH 25

RESULT 13

US-08-484-211C-41
Sequence 41, Application US/08484211C
Patent No. 5972645
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary L.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,211C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-211C-41

Query Match
Best Local Similarity 75.0%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 SODYFH 7
:|||||
DB 20 AEDYFH 25

RESULT 14

US-08-906-769-41
Sequence 41, Application US/08906769
Patent No. 6077687
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,769
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-906-769-41

Query Match
Best Local Similarity 75.0%; Score 30; DB 3; Length 25;
Best Local Similarity 66.7%; Pred. No. 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 SODYFH 7
:|||||
DB 20 AEDYFH 25

RESULT 15

US-08-906-616-41
Sequence 41, Application US/08906616
Patent No. 6121035
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 190

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; City: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,616
; FILING DATE: 05-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-906-616-41

Query Match 75.0%; Score 30; DB 3; Length 25;
Best Local Similarity 66.7%; Pred No 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 SQDYFH 7
: : : : :
Db 20 AEDYFH 25

Search completed: June 13, 2002, 08:54:31
Job time: 316 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:17 ; Search time 125.81 Seconds

(without alignments)
11.000 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45

Sequence: 1 DGNVPVPM 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	100.0	499	3	Q96TH4	Q96TH4 aspergillus
2	37	82.2	6260	2	Q54299	Q54299 streptomyc
3	36	80.0	335	4	Q9BWN2	Q9BWN2 homo sapien
4	36	80.0	360	4	Q9BQ81	Q9BQ81 homo sapien
5	36	80.0	457	16	Q987U9	Q987U9 rhizobium 1
6	36	80.0	619	5	Q9VKJ8	Q9VKJ8 drosophila
7	36	80.0	1385	12	Q65565	Q65565 bovine heip
8	35	77.8	98	5	Q9U684	Q9U684 echinometra
9	35	77.8	612	10	Q81829	Q81829 arabidopsis
10	35	77.8	650	10	Q94209	Q94209 oryza sativ
11	35	77.8	652	12	Q10436	Q10436 mouse adeno
12	35	77.8	722	2	P96091	P96091 treponema d
13	35	77.8	1392	12	Q9ELX9	Q9ELX9 cercopithe
14	35	77.8	1989	11	Q9ERT7	Q9ERT7 mus musculu
15	35	77.8	8563	2	Q54297	Q54297 streptomyc
16	35	77.8	10223	2	Q54296	Q54296 streptomyc

17	34	75.6	297	16	Q928V2	Q928V2 chlamydia p
18	34	75.6	335	16	Q98KV6	Q98KV6 rhizobium 1
19	34	75.6	465	2	Q69840	Q69840 streptomyc
20	34	75.6	472	11	Q9DAP3	Q9DAP3 mus musculu
21	34	75.6	585	5	Q9VJ36	Q9VJ36 drosophila
22	34	75.6	593	10	Q93VE2	Q93VE2 oryza sativ
23	34	75.6	640	3	Q13296	Q13296 aspergillus
24	34	75.6	732	5	Q9GR74	Q9GR74 leishmania
25	34	75.6	989	16	Q9CNT9	Q9CNT9 pasteurella
26	34	75.6	1390	10	Q49145	Q49145 arabidopsis
27	33.5	74.4	90	16	Q31786	Q31786 bacillus su
28	33	73.3	111	13	Q9DDT3	Q9DDT3 atalinga mi
29	33	73.3	111	13	Q9DDT1	Q9DDT1 ara araraun
30	33	73.3	125	17	Q59155	Q59155 pyrococcus
31	33	73.3	238	2	Q91IA9	Q91IA9 streptomyc
32	33	73.3	261	5	Q9XUR6	Q9XUR6 caenorhabdi
33	33	73.3	271	17	Q9HQT3	Q9HQT3 thermoplasm
34	33	73.3	288	17	Q96XP7	Q96XP7 sulfolobus
35	33	73.3	362	11	Q9CX00	Q9CX00 mus musculu
36	33	73.3	372	3	Q91BH2	Q91BH2 tricholoma
37	33	73.3	373	10	Q9FUC4	Q9FUC4 nicotiana t
38	33	73.3	403	16	Q9KAV1	Q9KAV1 bacillus ha
39	33	73.3	451	2	Q05074	Q05074 streptomyc
40	33	73.3	507	17	Q52024	Q52024 halobacteri
41	33	73.3	525	5	Q97229	Q97229 plasmodium
42	33	73.3	529	5	Q96KX9	Q96KX9 plasmodium
43	33	73.3	607	4	Q9NVM2	Q9NVM2 homo sapien
44	33	73.3	637	16	Q985R6	Q985R6 rhizobium 1
45	33	73.3	644	16	Q985R7	Q985R7 rhizobium 1

ALIGNMENTS

RESULT 1

Q96TH4 ID Q96TH4 PRELIMINARY: PRT; 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TRKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
GN AMYA OR AMY1.
OS Aspergillus oryzae, and
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062, 5059;
RP SEQUENCE FROM N.A.
RC SPECIES=A.oryzae; STRAIN=RTB40;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
RA Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator
RT gene, amyR, involved in the amylolytic gene expression in Aspergillus
RT oryzae.";
RT Biosci. Biotechnol. Biochem. 64:816-827(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=A.flavus; STRAIN=86-10D;
RA Pakhoury A.M., Woloshuk C.P.;
RT "Amyr, the alpha-amylase gene of Aspergillus flavus: Involvement in
RT aflatoxin biosynthesis in maize kernels.";
RL Phytopathology 89:908-914(1999).
DR EMBL; AF021876; BAA95703.1; -;
DR EMBL; AF139925; AAF14264.1; -;
SQ SEQUENCE 499 AA; 54804 MW; EEFA2ADA7ID20DA9 CRC64;

Query Match 100.0%; Score 45; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
 |||||
 DB 469 DGNVPVPM 476

RESULT 2
 ID 054299 PRELIMINARY; PRT: 6260 AA.
 AC 054299;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE POLYKETIDE SYNTHASE.
 GN RAPC.
 OS Streptomyces hygroscopicus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1912;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NRRL 5491;
 RC MEDLINE=95372374; PubMed=7644502;
 RA Schwecke T., Aparicio J.F., Molnar I., Koenig A., Khaw L.E.,
 Haydock S.F., O'Leary M., Caffrey P., Cortes J., Lester J.B.,
 Boehm G.A., Staunton J., Leadlay P.F.;
 RT "The biosynthetic gene cluster for the polyketide immunosuppressant
 rapamycin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:7839-7843(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NRRL 5491;
 RC MEDLINE=96186895; PubMed=8635730;
 RA Molnar I., Aparicio J.F., Haydock S.F., Khaw L.E., Schwecke T.,
 Koenig A., Staunton J., Leadlay P.F.;
 RT "Organisation of the biosynthetic gene cluster for rapamycin in
 Streptomyces hygroscopicus: analysis of genes flanking the polyketide
 synthase.";
 RL Gene 169:1-7(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NRRL 5491;
 RC MEDLINE=96186896; PubMed=8635756;
 RA Aparicio J.F., Molnar I., Schwecke T., Koenig A., Haydock S.F.,
 Khaw L.E., Staunton J., Leadlay P.F., Staunton J., Leadlay P.F.;
 RT "Organisation of the biosynthetic gene cluster for rapamycin in
 Streptomyces hygroscopicus: analysis of the enzymatic domains in the
 modular polyketide synthase.";
 RL Gene 169:9-16(1996).
 DR EMBL: X86780; CAA60462.1; -.
 DR HSSP: P25715; 1MLA.
 DR InterPro: IPR001227; Acyltransf_domain.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR000794; Ketoacyl_synth.
 DR InterPro: IPR003880; Phosphopant_attach.
 DR InterPro: IPR002364; QOR_zeta_crystal.
 DR Pfam: PF00698; Acyl_transf. 4.
 DR Pfam: PF00107; adh_zinc. 1.
 DR Pfam: PF00109; ketoacyl_synth. 4.
 DR Pfam: PF02801; ketoacyl_synth_C. 4.
 DR Pfam: PF00550; pp-binding. 4.
 DR PROSITE: PS50075; ACP_DOMAIN. 4.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE. 4.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEININE; 4.
 DR PROSITE: PS01162; QOR_ZETA_CRYSTAL; 1.
 KM Phosphopantetheine; Transferase.
 SO SEQUENCE 6260 AA; 659509 MW; D7D6C002F91F2080 CRC64;

Query Match 82.2%; Score 37; DB 2; Length 6260;
 Best Local Similarity 85.7%; Pred. No. 7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DGNVPVPM 7
 |||||
 DB 4372 DGNVPVPM 4378

RESULT 3
 ID 09BMN2 PRELIMINARY; PRT: 335 AA.
 AC 09BMN2;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE SIMILAR TO KIAA0174 GENE PRODUCT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EYE, RETINOBLASTOMA;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC000116; AAH00116.1; -.
 SO SEQUENCE 335 AA; 36622 MW; 7D0714269380E835 CRC64;

Query Match 80.0%; Score 36; DB 4; Length 335;
 Best Local Similarity 75.0%; Pred. No. 41;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
 |||||
 DB 225 DGTVPVPM 232

RESULT 4
 ID 09BQ81 PRELIMINARY; PRT: 360 AA.
 AC 09BQ81;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE SIMILAR TO KIAA0174 GENE PRODUCT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LUNG CARCINOMA;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LUNG CARCINOMA;
 RC TISSUE-LUNG CARCINOMA;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC000435; AAH04359.1; -.
 DR EMBL: BC000430; AAH00430.1; -.
 SO SEQUENCE 360 AA; 39927 MW; 5D0DD8A113DAA2BC CRC64;

Query Match 80.0%; Score 36; DB 4; Length 360;
 Best Local Similarity 75.0%; Pred. No. 45;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
 |||||
 DB 225 DGTVPVPM 232

```

RESULT      5
Q98709      PRELIMINARY;      PRT;      457 AA.
AC Q98709;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE GAMMA-Glutamylcysteine synthetase.
GN MLI6902.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Pseudomonadaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAF30309;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003010; BAB53101.1; -.
KW Complete proteome.
SQ SEQUENCE 457 AA; 51437 MW; 9DFE7989DD10CF38 CRC64;

```

```

Query Match      80.0%; Score 36; DB 16; Length 457;
Best Local Similarity 85.7%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 DGNPVP 7
   111 111
DB 45 DGNAPVP 51

```

```

RESULT      6
Q9VKJ8      PRELIMINARY;      PRT;      619 AA.
ID Q9VKJ8;
AC Q9VKJ8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG6201 PROTEIN.
GN CG6201.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRKLEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

```

```

RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mourtlov G., Mlshina N.V., Modary C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleby J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003631; AAF53070.1; -.
DR FLYBase: FBgn0032343; CG6201.
DR InterPro: IPR000514; Glyco_hydro.39.
DR Pfam: PF01229; Glyco_hydro.39; 1.
DR PRINTS: PR00745; GHYDRASE39.
SQ SEQUENCE 619 AA; 70341 MW; A7A79C7BF93A14DA CRC64;

```

```

Query Match      80.0%; Score 36; DB 5; Length 619;
Best Local Similarity 75.0%; Pred. No. 83;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 DGNVPPM 8
   111 111
DB 207 DGNVPLPM 214

```

```

RESULT      7
Q65565      PRELIMINARY;      PRT;      1385 AA.
ID Q65565;
AC Q65565;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MAJOR CAPSID PROTEIN.
GN ULI9.
OS Bovine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COOPER, AND JURA;
RX MEDLINE=9531343; PubMed=7793062;
RA Vilek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA Lechworth G.J., Schwytzer M.;
RT "Nucleotide sequence analysis of a 30-kb region of the bovine
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT the UL21 to UL4 genes of herpes simplex virus."
RL Virology 210:100-108(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JURA;
RX MEDLINE=9531343; PubMed=7793062;
RA Lowery D.E., Simard C., Bello L.J., Thiry E., Vilek C.;
RN Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN=JURA;
RX Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; Z48053; CA88114.1; -.
 DR EMBL; AJ004801; CA06115.1; -.
 DR InterPro; IPR000912; Herpes_MCP.
 DR Pfam; PF03122; Herpes_MCP.1.
 DR PRINTS; PR00235; HSVCAPSIDMCP.
 SO SEQUENCE 1385 AA; 149735 MW; 1B2117B7877060A3 CRC64;

Query Match 80.0%; Score 36; DB 12; Length 1385;
 Best Local Similarity 62.5%; Pred. No. 2e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DGNVPVP 8
 Db 580 GNTVPPL 587

RESULT 8
 O9U6E4 PRELIMINARY; PRT; 98 AA.
 AC O9U6E4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE BINDIN (FRAGMENT).
 OS Echinometra mathaei.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoida; Euechinoidea; Echinacea; Echinoida; Echinometridae;
 OC Echinometra.
 OK NCBI_TaxID=31178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M6B.1;
 RX MEDLINE=20006286; PubMed=10535974;
 RA Palumbi S.R.;
 RT "All males are not created equal: fertility difference depend on
 gamete recognition polymorphisms in sea urchins";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:12632-12637(1999).
 DR EMBL; AF186336; AAF07150.1; -.
 DR InterPro; IPR000775; Bindin.
 DR Pfam; PF02084; Bindin; 1.
 RT NON_TER 1
 FT NON_TER 98
 SO SEQUENCE 98 AA; 9660 MW; 579B07577FB70EFC CRC64;

Query Match 77.8%; Score 35; DB 5; Length 98;
 Best Local Similarity 85.7%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DGNVPVP 7
 Db 14 DGNVPVP 20

RESULT 9
 O81829 PRELIMINARY; PRT; 612 AA.
 AC O81829;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE GH3 LIKE PROTEIN.
 GN M4122.70 OR ATAG27260.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Reichert B.J., Barel E., Hohnsels J., Mewes H.W., Mayer K.,
 RA Scheller C., Bevan M.;

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Reichert B.J., Barel E., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL030978; CA19720.1; -.
 DR EMBL; AL161566; CAB79581.1; -.
 SO SEQUENCE 612 AA; 69283 MW; 8E0F91FE7BE2725A CRC64;

Query Match 77.8%; Score 35; DB 10; Length 612;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DGNVPVP 7
 Db 506 DGNVPVP 512

RESULT 10
 O94209 PRELIMINARY; PRT; 650 AA.
 AC O94209;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE POTASSIUM TRANSPORT PROTEIN.
 GN P0506E04.23.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OK NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0506E04";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003272; BAB67845.1; -.
 SO SEQUENCE 650 AA; 70963 MW; 9BDBC4F6870B4B50 CRC64;

Query Match 77.8%; Score 35; DB 10; Length 650;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GNPVPVP 8
 Db 160 GNPVPVP 166

RESULT 11
 ID 010436 PRELIMINARY; PRT; 652 AA.
 AC 010436;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PTP (FRAGMENT).
 GN E2B PTP.
 OS Mouse adenovirus type 1 (MAV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OK NCBI_TaxID=10530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Meissner J.D., Hirsch G.N., Larue E.A., Fulcher R.A., Spindler K.R.;
 RT "Completion of the DNA sequence of mouse adenovirus type 1: sequence

RT of E2B, I1, and I2 (18-51 map units)";
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U95843; AAB53751.1; -;
 DR InterPro: IPR003391; Adeno_terminal.
 DR Pfam: PF02459; Adeno_terminal; 1.
 FT NON_TER 1 1
 FT VARANT 644 644 E -> G.
 SQ SEQUENCE 652 AA; 74664 MW; 1400CF2D2E266D0B CRC64;

Query Match 77.8%; Score 35; DB 12; Length 652;
 Best Local Similarity 75.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 ||| |||
 Db 202 DGNVPVP 209

RESULT 12
 P96091 PRELIMINARY; PRT; 722 AA.
 AC P96091;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SERINE PROTEASE (DENTILISIN) FROM TREPONEMA DENTICOLA PRECURSOR.
 GN PRTP.
 OS Treponema denticola.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_TaxID=158;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=97101036; PubMed=8945563;
 RA Ishihara K., Miura T., Kuramitsu H.K., Oda K.;
 RT "Characterization of the Treponema denticola prtp gene encoding a
 RT prolyl-phenylalanine-specific protease (dentilisins).";
 RL Infect. Immun. 64:5178-5186(1996).
 DR EMBL: D83264; BAA11874.1; -;
 DR HSSP: Q45670; IDBI.
 DR MEROPS: S08.024; -;
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 2.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Signal; Protease.
 FT SIGNAL 1 26 POTENTIAL.
 SQ SEQUENCE 722 AA; 77474 MW; A604DA078846910D CRC64;

Query Match 77.8%; Score 35; DB 2; Length 722;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVP 7
 ||| |||
 Db 511 DGNVPVP 517

RESULT 13
 Q9EIX9 PRELIMINARY; PRT; 1392 AA.
 ID Q9EIX9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MAJOR CAPSID PROTEIN.
 OS Cercopithecine herpesvirus 7.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 RN [1]

OX NCBI_TaxID=35245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
 RT "Complete Sequence of the Simian Varicella Virus genome.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF275348; AAG27215.1; -;
 DR InterPro: IPR000912; Herpes_MCP.
 DR PRINTS: PR00235; HSVCAPSIDMCP.
 SQ SEQUENCE 1392 AA; 155857 MW; C94599418AF62E19 CRC64;

Query Match 77.8%; Score 35; DB 12; Length 1392;
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 .|||:|:|:
 Db 591 GNIPIPL 598

RESULT 14
 Q9ERT7 PRELIMINARY; PRT; 1989 AA.
 AC Q9ERT7;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SYNE-1B (FRAGMENT).
 GN SYNE1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=20493380; PubMed=10878022;
 RA Apel E.D., Lewis R.M., Grady R.M., Sanes J.R.;
 RT "Syne-1, a dystrophin- and klarsicht-related protein associated with
 RT synaptic nuclei at the neuromuscular junction.";
 RL J. Biol. Chem. 275:31986-31995(2000).
 DR EMBL: AF281870; AAG24393.1; -;
 DR MGD: MGI:1927152; Syne1.
 DR InterPro: IPR002017; Spectrin.
 DR Pfam: PF00435; Spectrin; 14.
 DR SMART: SM00150; SPEC; 14.
 FT NON_TER 1
 SQ SEQUENCE 1989 AA; 229066 MW; 381E6AF211CC138D CRC64;

Query Match 77.8%; Score 35; DB 11; Length 1989;
 Best Local Similarity 71.4%; Pred. No. 4.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 GNVVPVP 8
 ||:|:|:
 Db 786 GNIPIPL 792

RESULT 15
 Q54297 PRELIMINARY; PRT; 8563 AA.
 ID Q54297;
 AC Q54297;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE POLYKETIDE SYNTHASE.
 GN RAPA.
 OS Streptomyces hygroscopicus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1912;
 RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-NRRL 5491;
RX MEDLINE=95372374; PubMed=7644502;
RA Schwecke T., Aparicio J.F., Molnar I., Koenig A., Khaw L.E.,
RA Haydock S.F., Oillynk M., Caffrey P., Cortes J., Lester J.B.,
RA Boehm G.A., Staunton J., Leadlay P.F.;
RT "The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:7839-7843(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NRRL 5491;
RX MEDLINE=96186895; PubMed=8635730;
RA Molnar I., Aparicio J.F., Haydock S.F., Khaw L.E., Schwecke T.,
RA Koenig A., Staunton J., Leadlay P.F.;
RT "Organisation of the biosynthetic gene cluster for rapamycin in Streptomyces hygroscopicus: analysis of genes flanking the polyketide synthase.";
RL Gene 169:1-7(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-NRRL 5491;
RX MEDLINE=96186896; PubMed=8635756;
RA Aparicio J.F., Molnar I., Schwecke T., Koenig A., Haydock S.F.,
RA Khaw L.E., Staunton J., Leadlay P.F., Staunton J., Leadlay P.F.;
RT "Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hygroscopicus: analysis of the enzymatic domains in the modular polyketide synthase.";
RL Gene 169:9-16(1996).
DR EMBL: X86780; CAA60460.1; -.
DR HSSP: P08659; ILCI.
DR InterPro: IPR001227; Acyltransf. domain.
DR InterPro: IPR002085; Adh. zn. family.
DR InterPro: IPR000873; AMP-bind.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003880; Phosphopant. attach.
DR InterPro: IPR002364; OOR_zeta_crystal.
DR Pfam: PF00698; Acyl-transf. 4.
DR Pfam: PF00107; adh_zinc; 3.
DR Pfam: PF00501; AMP-binding; 1.
DR Pfam: PF00109; ketoacyl-synt; 4.
DR Pfam: PF02801; ketoacyl-synt_C; 4.
DR Pfam: PF00550; pp-binding; 5.
DR PRINTS: PR00154; AMB BINDING.
DR PROSITE: PS50075; ACP_DOMAIN; 5.
DR PROSITE: PS00455; AMP_BINDING; 1.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 4.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_3.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 5.
DR PROSITE: PS01162; OOR_ZETA_CRYSTAL; 2.
DR Phosphatetheine; transferase.
SQ SEQUENCE 8563 AA; 899948 MW; FEA462CD33838A37 CRC64;

Query Match 77.88; Score 35; DB 2; Length 8563;
Best Local Similarity 85.78; Pred. NO. 2.5e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 DGNVPP 7
DB 4389 DGDVPP 4395

Search completed: June 13, 2002, 09:11:19
Job time: 1064 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:33 ; Search time 57.95 Seconds
(without alignments)
3.372 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45

Sequence: 1 DGNVPVPW 8

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/prodata/2/1aa/CTUS.COMB.pep: *
6: /cgn2_6/prodata/2/1aa/Dackflle1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	478	1	US-08-720-899-7
2	45	100.0	478	2	US-08-459-610-7
3	45	100.0	478	2	US-08-343-804-7
4	45	100.0	478	2	US-08-339-715A-2
5	45	100.0	478	2	US-08-600-908A-10
6	45	100.0	478	3	US-08-683-838A-10
7	45	100.0	478	4	US-09-182-859-7
8	32	71.1	14	5	US-08-408-604A-15
9	32	71.1	14	5	PCT-US93-09626-15
10	32	71.1	230	3	US-09-167-434-2
11	32	71.1	230	3	US-08-853-755-2
12	31	68.9	261	4	US-09-045-973-3
13	30	66.7	40	2	US-08-448-481-4
14	30	66.7	45	1	US-08-377-687-41
15	30	66.7	45	2	US-08-777-192-41
16	30	66.7	45	4	US-08-971-982-41
17	30	66.7	53	1	US-08-640-847C-40
18	30	66.7	94	1	US-08-591-498-4
19	30	66.7	112	1	US-08-326-117B-17
20	30	66.7	112	3	US-08-982-129-17
21	30	66.7	338	4	US-09-413-231-6
22	30	66.7	338	4	US-09-413-231-7
23	30	66.7	449	5	US-08-657-392-2
24	30	66.7	449	5	PCT-US94-02539-2
25	30	66.7	459	2	US-08-673-312-2
26	30	66.7	471	2	US-08-657-392-31
27	30	66.7	471	5	PCT-US94-02539-31

28	30	66.7	589	2	US-08-756-317-5	Sequence 5, Appl1
29	30	66.7	711	1	US-08-235-838-7	Sequence 7, Appl1
30	30	66.7	711	2	US-08-465-473B-7	Sequence 7, Appl1
31	30	66.7	893	4	US-09-514-302-4	Sequence 4, Appl1
32	30	66.7	907	4	US-08-989-299-12	Sequence 12, Appl1
33	30	66.7	1253	1	US-07-920-281C-3	Sequence 3, Appl1
34	30	66.7	1253	4	US-08-466-277-3	Sequence 3, Appl1
35	30	66.7	1528	1	US-08-326-117B-2	Sequence 2, Appl1
36	30	66.7	1528	3	US-08-982-129-2	Sequence 2, Appl1
37	30	66.7	1938	4	US-09-514-302-2	Sequence 2, Appl1
38	29	64.4	33	4	US-08-810-712-30	Sequence 30, Appl1
39	29	64.4	337	1	US-08-440-856A-3	Sequence 12, Appl1
40	29	64.4	415	1	US-08-451-715A-12	Patent No. 5194375
41	29	64.4	459	6	5194375-6	Sequence 2, Appl1
42	29	64.4	524	1	US-08-461-837-2	Sequence 12, Appl1
43	29	64.4	524	2	US-08-480-736-12	Sequence 12, Appl1
44	29	64.4	524	3	US-08-973-223-2	Sequence 2, Appl1
45	29	64.4	524	4	US-09-347-060-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thøgersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: ARTILASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720, 899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343, 804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr. Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-720-899-7

Query Match 100.0%; Score 45; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.4; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

OY 1 DGNVPPM 8
 |||||
 Db 448 DGNVPPM 455

RESULT 2

US-08-459-610-7

Sequence 7, Application US/08459610

Patent No. 5801043

GENERAL INFORMATION:

APPLICANT: Bisgaard-Frantzen, Henrik

APPLICANT: Borchert, Torben Vedel

APPLICANT: Svendsen, Allan

APPLICANT: Thellersen, Marianne

TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESS: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10174-6401

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,610

FILING DATE: 02-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/343,804

FILING DATE: 22-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lowney Dr., Karen A.

REGISTRATION NUMBER: 31,274

REFERENCE/DOCKET NUMBER: 4054, 214-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 478 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-459-610-7

Query Match 100.0%; Score 45; DB 1; Length 478;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPPM 8
 |||||
 Db 448 DGNVPPM 455

RESULT 3

US-08-343-804-7

Sequence 7, Application US/08343804

Patent No. 5830837

GENERAL INFORMATION:

APPLICANT: Bisgaard-Frantzen, Henrik

APPLICANT: Borchert, Torben Vedel

APPLICANT: Svendsen, Allan

APPLICANT: Thellersen, Marianne

TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESS: No. 58308370 No. 5830837disk of No. 5830837th America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10174-6401

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/343,804

FILING DATE: 22-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Lowney Dr., Karen A.

REGISTRATION NUMBER: 31,274

REFERENCE/DOCKET NUMBER: 4054, 214-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 478 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-343-804-7

Query Match 100.0%; Score 45; DB 2; Length 478;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPPM 8
 |||||
 Db 448 DGNVPPM 455

RESULT 4

US-08-339-715A-2

Sequence 2, Application US/08339715A

Patent No. 5965442

GENERAL INFORMATION:

APPLICANT: Kaneko, Hiroki

APPLICANT: Takada, Toshikazu

APPLICANT: Shimada, Jiro

APPLICANT: Kuriki, Takashi

APPLICANT: Yanase, Michiyo

APPLICANT: Takata, Hiroki

APPLICANT: Okada, Shigetaka

TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESS: POPHAM, HALK, SCHNOBIRCH & KAUFMAN, LTD.

STREET: Metropolitan Square Building, Suite 800, 1450

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/339,715A

FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Melse, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 45; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8
|||||
DB 447 DGNVPVPM 454

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-600-908A-10

Query Match 100.0%; Score 45; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8
|||||
DB 448 DGNVPVPM 455

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-683-838A-10

Query Match 100.0%; Score 45; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8
|||||
DB 448 DGNVPVPM 455

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgard-Frantzen, Henrik

;; TITLE OF INVENTION: Alpha-Amylase Mutants
;; FILE REFERENCE: 4796.204-US
;; CURRENT APPLICATION NUMBER: US/09/182,859
;; CURRENT FILING DATE: 1998-10-29
;; EARLIER APPLICATION NUMBER: 0515/96
;; EARLIER FILING DATE: 1996-04-30
;; EARLIER APPLICATION NUMBER: 0712/96
;; EARLIER FILING DATE: 1996-06-28
;; EARLIER APPLICATION NUMBER: 0775/96
;; EARLIER FILING DATE: 1996-07-11
;; EARLIER APPLICATION NUMBER: 1263/96
;; EARLIER FILING DATE: 1996-11-08
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 7
;; LENGTH: 478
;; TYPE: PRT
;; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 45; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 1,4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 448 DGNVPVPM 455

RESULT 8
US-08-408-604A-15
; Sequence 15, Application US/08408604A
; Patent No. 5801149
; GENERAL INFORMATION:
; APPLICANT: Shoelson, Steven
; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/408,604A
; FILING DATE: 21-MAR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/134,558
; FILING DATE: 08-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/959,949
; FILING DATE: 09-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/722,359
; FILING DATE: 19-JUNE-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-014CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14

;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-408-604A-15

Query Match 71.1%; Score 32; DB 1; Length 14;
Best Local Similarity 62.5%; Pred. No. 6;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 3 DGNVPVPM 10

RESULT 9
PCT-US93-09626-15
; Sequence 15, Application PC/TUS9309626
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES
; NUMBER OF SEQUENCES: 48
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/09626
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/959,949
; FILING DATE: October 9, 1992
; APPLICATION NUMBER: 07/722,359
; FILING DATE: June 19, 1991
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US93-09626-15

Query Match 71.1%; Score 32; DB 5; Length 14;
Best Local Similarity 62.5%; Pred. No. 6;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 3 DGNVPVPM 10

RESULT 10
US-09-167-434-2
; Sequence 2, Application US/09167434
; Patent No. 6008033
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiyang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; TITLE OF INVENTION: No. 6008033tel Proteases, Compositions Capable of
; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA

```

;
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,434
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/853,755
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,470
; FILING DATE: 26-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Diner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: P50472-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5022
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-167-434-2

Query Match 71.1%; Score 32; DB 3; Length 230;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GNPVPM 8
Db 36 GNIPPL 42

RESULT 11
; US-08-853-755-2
; Sequence 2, Application US/08853755
; Patent No. 6083711
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiyang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6083711el Proteases, Compositions Capable of
; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
```

```

;
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,755
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,616
; FILING DATE: 15-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,470
; FILING DATE: 26-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Diner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: P50472-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5022
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-853-755-2

Query Match 71.1%; Score 32; DB 3; Length 230;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GNPVPM 8
Db 36 GNIPPL 42

RESULT 12
; US-09-045-973-3
; Sequence 3, Application US/09045973
; Patent No. 6165767
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Maiah
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
```

```

; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,973
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0491 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
;
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 261 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT18
; CLONE: 2534680
;
; US-09-045-973-3
;
Query Match      68.9%; Score 31; DB 4; Length 261;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNVFP 7
   11111
DB 70 DGAVIDP 76

RESULT 13
US-08-448-481-4
; Sequence 4, Application US/08448481
; Patent No. 5914270
; GENERAL INFORMATION:
; APPLICANT: Coulos-Thyenot, Pierre M.L.
; APPLICANT: Jouenne, Thierry G.
; APPLICANT: Maes, Olivier C.A.
; APPLICANT: Deloire, Alain J.
; APPLICANT: Boulay, Michel P.H.
; APPLICANT: Guern, Jean R.D.
; TITLE OF INVENTION: Method for Promoting the Differentiation
; TITLE OF INVENTION: of Plant Cells in Culture
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pollock, Vande Sande & Priddy
; STREET: 1990 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-3425
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,481
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 800
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01239
; FILING DATE: 14-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92 15044
; FILING DATE: 14-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Franklin, Eric J.
; REGISTRATION NUMBER: 37,134
; REFERENCE/DOCKET NUMBER: 0061/00065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 331-7111
; TELEFAX: (202) 293-6229
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
;
; US-08-448-481-4
;
Query Match      66.7%; Score 30; DB 2; Length 40;
Best Local Similarity 83.3%; Pred. No. 42;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVPVPM 8
   11111
DB 23 NVPVPL 28

RESULT 14
US-08-377-687-41
; Sequence 41, Application US/08377687
; Patent No. 5538525
; GENERAL INFORMATION:
; APPLICANT: BROEKERT, WILLEM F.
; APPLICANT: CAMME, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZER
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,687
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-862-0944
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
```


LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-377-687-41

Search completed: June 13, 2002, 08:54:34
Job time: 319 sec

Query Match 66.7%; Score 30; DB 1; Length 45;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 NVPVPM 8
|||||:
DB 23 NVPVPL 28

RESULT 15
US-08-777-192-41
Sequence 41 Application US/08777192
Patent No. 5824869
GENERAL INFORMATION:
APPLICANT: BROEKERT, WILLEM F.
APPLICANT: CAMMUE, BRUNO P.A.
APPLICANT: OSBORN, RUPERT W.
APPLICANT: REES, SARAH B.
APPLICANT: TERRAS, FRANKY R.G.
APPLICANT: VANDERLEIJEN, JOZEF
TITLE OF INVENTION: BIOCIDAL PROTEINS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DABRY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/777,192
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/002,480
FILING DATE: 04-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKUTIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 99042/SEE. 36525/US/A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-777-192-41

Query Match 66.7%; Score 30; DB 2; Length 45;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 NVPVPM 8
|||||:
DB 23 NVPVPL 28

Fri Jun 14 10:36:51 2002

us-09-710-339-2_copy_468_475.rai

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:55:56 ; Search time 73.31 seconds

(without alignments)
10.486 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45
Sequence: 1 DGNVPVPM 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 28338 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 28338

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR-71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	478	2 JK0201	alpha-amylose (EC
2	45	100.0	498	2 A48305	alpha-amylose (EC
3	45	100.0	499	1 ALAS1	alpha-amylose (EC
4	45	100.0	499	1 ALAS1	alpha-amylose (EC
5	45	100.0	499	2 B48305	alpha-amylose (EC
6	45	100.0	499	2 JS0663	alpha-amylose (EC
7	45	100.0	499	2 JT0466	alpha-amylose (EC
8	45	100.0	499	2 JN0588	alpha-amylose (EC
9	37	82.2	6260	2 T30228	polyketide synthas
10	36	80.0	336	2 AG0511	diolp-glucose 4,6-d
11	36	80.0	428	2 AG0511	FIXC protein limpo
12	36	80.0	446	2 B70776	probable glp2 - My
13	36	80.0	837	1 A31842	endo-1,4-beta-xyla
14	36	80.0	1385	2 S61236	major capsid prote
15	36	80.0	1396	1 VCB40	auxin-regulated pr
16	35	77.8	612	2 T05750	polyketide synthas
17	35	77.8	8563	2 T30226	polyketide synthas
18	35	77.8	10223	2 T30225	5'-methylthioaden
19	34	75.6	297	2 E72102	hypothetical prote
20	34	75.6	465	2 T34690	hypothetical prote
21	34	75.6	926	2 T36640	chitin synthase 2
22	34	75.6	1038	2 T13177	soy protein - Tru1
23	34	75.6	1390	2 T31353	hypothetical prote
24	34	75.6	1974	2 T16703	hypothetical prote
25	33.5	74.4	90	2 H69886	hypothetical prote
26	33	73.3	81	2 JN0058	hypothetical 8.9K
27	33	73.3	125	2 A71024	hypothetical prote
28	33	73.3	261	2 T24532	hypothetical prote
29	33	73.3	346	2 A13587	membrane dipeptid

30	33	73.3	379	2 A48936	nitrite reductase
31	33	73.3	403	2 C83924	aspartate aminotra
32	33	73.3	430	2 AC3037	hypothetical prote
33	33	73.3	451	2 S30401	hypothetical prote
34	33	73.3	471	2 H98248	nitrate-binding pr
35	33	73.3	484	1 A35282	alpha-amylose (EC
36	33	73.3	507	2 T08337	hypothetical prote
37	33	73.3	671	2 A97352	penicillin-binding
38	33	73.3	760	2 AB2225	hypothetical prote
39	33	73.3	870	2 A86567	DNA polymerase I
40	33	73.3	870	2 F72058	DNA polymerase I
41	33	73.3	1061	2 T13107	tail tip fiber pro
42	33	73.3	1375	2 T42586	major capsid prote
43	33	73.3	1376	1 VCBED6	major capsid prote
44	33	73.3	1484	2 T29275	hypothetical prote
45	33	71.1	149	2 C72410	actinorhodin polyk

ALIGNMENTS

RESULT 1
JK0201
Alpha-amylose (EC 3.2.1.1) - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amylose A
C:Species: Aspergillus oryzae
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JK0201
R:Toda, H.; Kondo, K.; Narita, K.
Proc. Jpn. Acad. 58B, 208-212, 1982
A>Title: The complete amino acid sequence of Taka-amylose A.
A:Reference number: JK0201
A:Accession: JK0201
A:Molecule type: protein
A:Residues: 1-478 <TOP>
C:Comment: One atom of calcium per molecule is essential for the activity.
C:Function: See also PIR:JT0466 and PIR:JS0240.
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:173-300/Domain: alpha-amylose core homology <AMT>
F:197/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:210/230/297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 45; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DGNVPVPM 8
Db 447 DGNVPVPM 454

RESULT 2
A48305
Alpha-amylose (EC 3.2.1.1) A precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmone, C.L.; Kodama, K.H.; Royer, T.J
Curr. Genet. 17, 203-212, 1990
A>Title: Cloning, characterization, and expression of two alpha-amylose genes from As
A:Accession number: A48305; WUID:90254827
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-498 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amyLase; alpha-amyLase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amyLase core homology <AMY>

Query Match 100.0%; Score 45; DB 2; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DGNVPPM 8
 Db 469 DGNVPPM 476

RESULT 3

ALAS1

alpha-amyLase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
 M:Alternate names: alpha-amyLase isozyme II; glycogenase; taka-amyLase A
 C:Species: Aspergillus oryzae
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2000
 C:Accession: S04548; A33214; J50240; A91930; A93767; A10627

R:Witsell, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amyLase genes of Aspergillus oryzae exhibit identical intron-exon
 A:Reference number: S04548; MUID:89237897

A:Accession: S04548

A:Molecule type: DNA

A:Residues: 1-499 <WR>

A:Cross-references: EMBL:X12725; NID:92430; PIDN:CAA31218.1; PID:9295921

A:Accession: A33214

A:Molecule type: mRNA

A:Residues: 1-499 <WR>

A:Cross-references: GB:X12725; NID:92430; PIDN:CAA31218.1; PID:9295921

R:Genes, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989

A:Title: Aspergillus oryzae has two nearly identical Taka-amyLase genes, each containing
 A:Reference number: J50240; MUID:89378767

A:Accession: J50240

A:Molecule type: DNA

A:Residues: 1-499 <GEN>

A:Genetics: AMT2

A>Note: The authors refer to this as isozyme II

R:Isemura, S.; Ikenaka, T.

J. Biochem. 74, 1-10, 1973

A:Reference number: A91930; MUID:74001521

A:Accession: A91930

A:Molecule type: protein

A:Residues: 206-225 <ISE>

R:Narita, K.

Proc. Jpn. Acad. 51, 285-290, 1975

A:Reference number: A93767

A:Accession: A93767

A:Molecule type: protein

A:Residues: 434-443,446-447, 'Q',449-458, 'GTYV',459-464,467-468, 'B',470, 'B',472-499 <NMR>

R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.

J. Biochem. 95, 697-702, 1984

A:Title: Structure and possible catalytic residues of Taka-amyLase A.

A:Reference number: A37454; MUID:84212370

A:Contents: annotation; X-ray crystallography, 3.0 angstroms

R:Swift, H.J.; Brady, L.; Deewenda, S.S.; Dodson, E.J.; Turkemburg, J.P.; Wilkinson, A.

submitted to the Brookhaven Protein Data Bank, August 1992

C:Comment: One atom of calcium per molecule is essential for activity.

C:Genetics: <AMY1>

A:Gene: amy1

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Genetics: <AMY2>

A:Gene: amy2; AmyII

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amyLase; alpha-amyLase core homology

C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po

F:1-21/Domain: signal sequence #status predicted <SIG>

F:194-321/Product: alpha-amyLase 1 #status experimental <MAT>

F:51-59,171-185,261-304,461-496/disulfide bonds: #status experimental

F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted

F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 45; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPPM 8
 Db 469 DGNVPPM 476

RESULT 4

ALAS3

alpha-amyLase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
 M:Alternate names: alpha-amyLase isozyme I; glycogenase; Taka-amyLase A
 C:Species: Aspergillus oryzae
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713

R:Witsell, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amyLase genes of Aspergillus oryzae exhibit identical intron-exo
 A:Reference number: S04548; MUID:89237897

A:Accession: S04549

A:Molecule type: DNA

A:Residues: 1-499 <WR>

A:Cross-references: EMBL:X12727; NID:92454; PIDN:CAA31220.1; PID:9295922

A:Accession: A33215

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-499 <WR>

A:Cross-references: GB:X12727; NID:92454; PIDN:CAA31220.1; PID:9295922

R:Genes, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989

A:Title: Aspergillus oryzae has two nearly identical Taka-amyLase genes, each contain
 A:Reference number: J50240; MUID:89378767

A:Accession: A44713

A:Molecule type: DNA

A:Residues: 1-499 <GEN>

A>Note: the authors refer to this as isozyme I

R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.

J. Biochem. 95, 697-702, 1984

A:Title: Structure and possible catalytic residues of Taka-amyLase A.

A:Reference number: A37454; MUID:84212370

A:Contents: annotation; X-ray crystallography, 3.0 angstroms

C:Comment: One atom of calcium per molecule is essential for activity.

C:Genetics:

A:Gene: amy1

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amyLase; alpha-amyLase core homology

C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amyLase 3 #status experimental <MAT>

F:194-321/Domain: alpha-amyLase core homology <AMY>

F:51-59,171-185,261-304,461-496/disulfide bonds: #status experimental

F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted

F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 45; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 |||||||
 Db 469 DGNVPVP 476

RESULT 5
 B48305
 alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
 C:Species: Aspergillus awamori
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
 C:Accession: B48305
 R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
 Curr. Genet. 17, 203-212, 1990
 A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Asper
 A:Reference number: A48305; MUID:90254827
 A:Accession: B48305
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1499 <KOR>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 C:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 45; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 |||||||
 Db 469 DGNVPVP 476

RESULT 6
 JS0663
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
 C:Species: Aspergillus sp.
 C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
 C:Accession: JS0663
 R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
 Biosci. Biotechnol. Biochem. 56, 174-179, 1992
 A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousanii and its expression
 A:Reference number: JS0663; MUID:92323146
 A:Accession: JS0663
 A:Molecule type: mRNA
 A:Residues: 1499 <SHI>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <AMP>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 45; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 |||||||
 Db 469 DGNVPVP 476

RESULT 7
 JT0466
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 N:Alternate names: Takamylase A
 C:Species: Aspergillus oryzae
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
 C:Accession: JT0466
 R:Tada, S.; Imura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
 Agric. Biol. Chem. 53, 593-599, 1989
 A:Title: Cloning and nucleotide sequence of the genomic Takamylase A gene of Asperg
 A:Reference number: JT0466
 A:Accession: JT0466
 A:Molecule type: DNA
 A:Residues: 1499 <TAD>
 C:Comment: See also PIR:JK0201 and PIR:JS0240.
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Gene(s):
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <AMY>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 45; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 |||||||
 Db 469 DGNVPVP 476

RESULT 8
 JN0588
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 N:Alternate names: Taka-amylase A
 C:Species: Aspergillus oryzae
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
 C:Accession: JN0588
 R:Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsudoi, A.; Udaoka, S.
 Gene 84, 319-327, 1989
 A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
 A:Reference number: JN0588; MUID:90128276
 A:Accession: JN0588
 A:Molecule type: mRNA
 A:Residues: 1499 <TSU>
 C:Comment: The alpha amylases are encoded by multigene family.
 C:Gene(s):
 A:Gene: Taa-G1
 A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <AMY>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 45; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8

Db 469 DGNVPVP 476
|||||

RESULT 9
T30228
polyketide synthase - Streptomyces hygroscopicus
C:Species: Streptomyces hygroscopicus
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
C:Accession: T30228
R:Appelton, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Straub
Gene 165, 9-16, 1996
A:Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hygroscopicus
A:Reference number: Z20782; MIMD:96186886
A:Accession: T30228
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6260 <APA>
A:Cross-references: EMBL:X86780; NID:q987088; PID:q987102; PIDN:CAA60462.1
C:Genetics:
A:Gene: rapC
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protein
C:Keywords: carrier protein
F:55-451/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <AS1>
F:1536-1607/Domain: acyl carrier protein homology <ACP1>
F:1651-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <AS2>
F:3123-3194/Domain: acyl carrier protein homology <ACP2>
F:3228-3632/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <AS3>
F:3725-3999/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F:5114-5185/Domain: acyl carrier protein homology <ACP3>
F:5242-5638/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <AS4>
F:5759-6044/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F:6135-6206/Domain: acyl carrier protein homology <ACP4>

Query Match 82.2%; Score 37; DB 2; Length 6260;
Best Local Similarity 85.7%; Pred. No. 4; 5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 7
|||||
Db 4372 DGNMPVP 4378

RESULT 10
AG3564
dcdp-glucose 4,6-dehydratase (EC 4.2.1.46) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AG3564
R:DeVeechio, V.G.; Kaputrat, V.; Redkar, R.J.; Patra, G.; Mufer, C.; Los, T.; Ivanova,
Mazur, M.; Goldsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Lelesg
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-536 <KUR>
A:Cross-references: GB:AE008918; PIDN:AL153682.1; PID:q17984602; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110440
A:Map position: 11
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 80.0%; Score 36; DB 2; Length 336;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 DGNVPVP 8
|||||

Db 186 DNNLPVP 193

RESULT 11
AG0511
fixC protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain C)
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG0511
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
S.; Moutie, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A:Reference number: AB0502; PMID:11677608
A:Accession: AG0511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <PAR>
A:Cross-references: GB:AL15382; PIDN:CAD01231.1; PID:q16501360; GSPDB:GN00176
C:Genetics:
A:Gene: STY0087
C:Superfamily: fixC protein

Query Match 80.0%; Score 36; DB 2; Length 428;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
|||||
Db 394 DGSAPVP 401

RESULT 12
B70776
probable gln2 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70776
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MIMD:98295987
A:Accession: B70776
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-446 <COL>
A:Cross-references: GB:T70692; GB:AL123456; NID:q3261567; PIDN:CAA94645.1; PID:q12619
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: glnA2
C:Superfamily: glutamate--ammonia ligase

Query Match 80.0%; Score 36; DB 2; Length 446;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
|||||
Db 144 DGSVPVP 151

RESULT 13
A31842
endo-1,4-beta-xylanase (EC 3.2.1.8) z precursor - Clostridium thermocellum
N:Alternate names: xylanase Z

C.Species: Clostridium thermocellum
 C.Date: 31-Mar-1990 #sequence_revision 11-Apr-1997 #text_change 18-Jun-1999
 C.Accession: A31842
 R.Grepinet, O.; Chebrou, M.C.; Beguin, P.
 J. Bacteriol. 170, 4582-4588, 1988
 A.Title: Nucleotide sequence and deletion analysis of the xylanase gene (xynZ) of Clostridium thermocellum
 A.Reference number: A31842; MUID:89008072
 A.Accession: A31842
 A.Molecule type: DNA
 A.Residues: 1-837 <GRE>
 A.Cross-references: GB:M2624; NID:g144931; PIDN:AAA23286.1; PID:g144932
 C.Genetics:
 A.Gene: xynZ
 C.Function:
 A.Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
 A.Pathway: xylan degradation
 C.Superfamily: Clostridium endo-1,4-beta-xylanase Z; Clostridium cellulase repeat homolog
 C.Keywords: duplication; extracellular protein; glycosidase; heat-stable protein; hydrolysis
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-837/Product: endo-1,4-beta-xylanase #status predicted <MAT>
 F:326-419/Domain: Clostridium xylanase A repeat homology <CXA>
 F:430-453/Domain: Clostridium cellulase repeat homology <CCR1>
 F:464-487/Domain: Clostridium cellulase repeat homology <CCR2>
 F:548-834/Domain: Streptomyces endo-1,4-beta-xylanase A homology <SXY>
 F:645,754/Active site: Glu #status predicted

Query Match 80.0%; Score 36; DB 1; Length 837;
 Best Local Similarity 85.7%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DGNVPVP 7
 ||| |||
 Db 284 DGNVPVP 290

RESULT 14
 S61236
 major capsid protein - bovine herpesvirus 1
 C.Species: bovine herpesvirus 1
 C.Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 26-Aug-1999
 C.Accession: S61236
 R.Vilek, C.; Benes, V.; Lu, Z.; Kutlish, G.F.; Paces, V.; Rock, D.; Letchworth, G.J.; Schmitt, J.
 submitted to the EMBL Data Library, January 1995
 A.Description: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome
 A.Reference number: S61233
 A.Accession: S61236
 A.Status: Preliminary
 A.Molecule type: DNA
 A.Residues: 1-1385 <VIC>
 A.Cross-references: EMBL:Z48053; NID:g971311; PIDN:CAA8114.1; PID:g971315
 C.Superfamily: varicella-zoster virus major capsid protein

Query Match 80.0%; Score 36; DB 2; Length 1385;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DGNVPVP 8
 :||:||||
 Db 580 NGNVPPL 587

RESULT 15
 VCB840
 major capsid protein - human herpesvirus 3
 C.Species: human herpesvirus 3; varicella-zoster virus
 C.Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
 C.Accession: E27341
 R.Davison, A.J.; Scott, J.E.
 J. Gen. Virol. 67, 1759-1816, 1986
 A.Title: The complete DNA sequence of varicella-zoster virus.
 A.Reference number: A27345; MUID:86306657

A.Accession: E27341
 A.Molecule type: DNA
 A.Residues: 1-1396 <DAV>
 A.Cross-references: EMBL:X04370; NID:g95989; PIDN:CAA27923.1; PID:g60029
 C.Genetics:
 A.Gene: 40
 C.Superfamily: varicella-zoster virus major capsid protein
 C.Keywords: capsid protein

Query Match 80.0%; Score 36; DB 1; Length 1396;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DGNVPVP 8
 :||:||||
 Db 595 NGNVPPL 602

Search completed: June 13, 2002, 08:55:58
 Job time: 378 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:55 ; Search time 38.1 Seconds
(without alignments)
8.130 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45

Sequence: 1 DGNVPVPM 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

105224

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	498	1	AMYA ASPAW
2	45	100.0	499	1	AMYA ASPOR
3	45	100.0	499	1	AMYA ASPAM
4	45	100.0	499	1	AMYA ASPSH
5	36	80.0	364	1	Y17A_HUMAN
6	36	80.0	446	1	GLN2_MYCTU
7	36	80.0	837	1	XYNZ_CLOTM
8	36	80.0	1396	1	VCAP_HSV6
9	34	75.6	926	1	CHS2_SCHPO
10	34	75.6	1038	1	SOG_DROME
11	34	75.6	1654	1	PCFB_HUMAN
12	33	73.3	125	1	RT06_MOUSE
13	33	73.3	379	1	NIR_PSESP
14	33	73.3	484	1	AMYA ASPNG
15	33	73.3	1376	1	VCAP_HSV6
16	32	71.1	380	1	GBR_ORYSA
17	32	71.1	528	1	VP40_HSV6U
18	32	71.1	678	1	CMC1_HUMAN
19	32	71.1	683	1	AMOH_ARTGO
20	32	71.1	1345	1	VCAP_HSV6U
21	32	71.1	1345	1	VCAP_HSV7J
22	31	68.9	124	1	RNP_PIG
23	31	68.9	131	1	Y1L2_STRAU
24	31	68.9	262	1	KK4_BACCT
25	31	68.9	327	1	FIPI_YEAST
26	31	68.9	397	1	TBXT_CHICK
27	31	68.9	514	1	UVS2_XENLA
28	31	68.9	521	1	PHLB_MYCTU
29	31	68.9	680	1	APB2_MOUSE
30	31	68.9	1066	1	HGPC_HAEIN
31	31	68.9	1067	1	HGPC_HAEIN
32	31	68.9	1109	1	MSH7_ARATH
33	31	68.9	1505	1	DPOA_DROME

ALIGNMENTS

RESULT	1	AMYA ASPAW	STANDARD:	PRT:	498 AA.
AC	002905;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase A).				
GN	AMTA.				
OS	Aspergillus awamori.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiaceae; Trichocomaceae; mitosporaceae; Aspergillius.				
OX	NCBI_TaxID=105351;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=UVR143F.				
RX	MEDLINE=90254627; PubMed=2340591;				
RA	Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,				
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,				
RA	Berka R.M.;				
RT	"Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus niger var. awamori.";				
RT	Curr. Genet. 17:203-212(1990).				
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides.				
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: X52755; CNA36966.1; -.				
DR	HSSP: P10529; 77RA.				
DR	InterPro: IPR00461; Alpha-amylase.				
DR	Plant: PP00128; alpha-amylase; 1.				
KW	Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;				
KW	Glycoprotein; Signal; Multigene family.				
FT	SIGNAL 1 21	POTENTIAL.			
FT	CHAIN 22 498	ALPHA-AMYLASE A.			
FT	ACT_SITE 227 227	BY SIMILARITY.			
FT	ACT_SITE 231 231	BY SIMILARITY.			
FT	ACT_SITE 251 251	BY SIMILARITY.			
FT	ACT_SITE 318 318	BY SIMILARITY.			
FT	DISULFID 51 59	BY SIMILARITY.			
FT	DISULFID 171 185	BY SIMILARITY.			
FT	DISULFID 261 304	BY SIMILARITY.			
FT	DISULFID 461 496	BY SIMILARITY.			
FT	CARBOHYD 218 218	N-LINKED (GLCNAC...) (POTENTIAL).			
SQ	SEQUENCE 498 AA: 54880 MW: 765851	IBCO1A8A01 CRC64;			

34	30	66.7	87	1	DB15_MOUSE	O09035 mus musculu
35	30	66.7	87	1	DB15_RAT	P56702 rattus norv
36	30	66.7	120	1	MTTP_DAUCA	P27631 daucus caro
37	30	66.7	171	1	Y708_TREPA	O83706 treponema p
38	30	66.7	189	1	DYR_CHICK	P00378 gallus gall
39	30	66.7	201	1	CYC4_PSEAE	P00106 pseudomonas
40	30	66.7	210	1	CYC4_PSEAE	Q52369 pseudomonas
41	30	66.7	241	1	UTR4_YEAST	P32626 saccharomyc
42	30	66.7	268	1	VG65_BPML5	O05279 mycobacteri
43	30	66.7	300	1	IPMS_ECOLI	P08191 escherichia
44	30	66.7	336	1	E13A_LYCES	O01412 lycopersico
45	30	66.7	338	1	IPNS_CEPAC	P03189 cephalospor

Query Match 100.0%; Score 45; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DGNVFPVM 8
 Db 469 DGNVFPVM 476

RESULT 2

AMTA_ASPOR STANDARD; PRT; 499 AA.
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 DE (1,4-alpha-D-glucan glucanohydrolase).
 GN AMT1 AND AMT2 AND AMT3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutriales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 63303;
 RA MEDLINE=89237897; PubMed=2785629;
 RA Witsell S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT intron-exon organization.";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89378767; PubMed=2789162;
 RA Genes M.J., Dove M.J., Selligy V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns.";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae.";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90128276; PubMed=2612911;
 RA Tsubagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes.";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A";
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RA MEDLINE=74001521; PubMed=4733850;
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RT A with trypsin and chymotrypsin.";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; PubMed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution.";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RA MEDLINE=84212370; PubMed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A.";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RA MEDLINE=97428212; PubMed=9283074;
 RA Brozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution.";
 RL Biochemistry 36:10837-10845(1997).
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I- COPFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I- MISCELLANEOUS: THE SEQUENCE OF AMT1 AND AMT2 IS SHOWN.
 CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sb.ch/announce/>
 CC or send an email to license@isb-sb.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04548; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: JT0466; JT0466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR GlycoSuiteDB: P10529;
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT CONFLICT 93 94
 FT CONFLICT 94 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTID-CAR.000125.
 O -> R (IN AMY3).
 F -> L (IN AMY3).
 FT -> DC (IN REF. 5).
 Q -> T (IN REF. 5).
 Q -> Y (IN REF. 3).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 G -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).

FT CONFLICT 370 370 L -> A (IN REF. 4).
 FT CONFLICT 406 409 MPIY -> PYI (IN REF. 5).
 FT CONFLICT 448 448 G -> S (IN REF. 5).
 FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).

FT CONFLICT 370 370
 FT CONFLICT 406 409
 FT CONFLICT 448 448
 FT CONFLICT 497 497
 FT CONFLICT 24 27
 FT TURN 28 29
 FT STRAND 32 35
 FT HELIX 37 40
 FT STRAND 42 42
 FT TURN 43 44
 FT HELIX 53 55
 FT STRAND 61 61
 FT HELIX 63 67
 FT TURN 68 69
 FT HELIX 70 74
 FT TURN 75 77
 FT STRAND 80 83
 FT STRAND 87 89
 FT STRAND 94 94
 FT TURN 95 96
 FT STRAND 97 97
 FT TURN 100 101
 FT STRAND 105 111
 FT TURN 113 115
 FT TURN 118 119
 FT HELIX 120 130
 FT TURN 131 132
 FT STRAND 134 139
 FT STRAND 143 143
 FT STRAND 146 146
 FT HELIX 150 152
 FT TURN 155 156
 FT STRAND 159 159
 FT HELIX 164 166
 FT STRAND 167 167
 FT STRAND 172 172
 FT HELIX 179 184
 FT STRAND 186 188
 FT STRAND 193 194
 FT STRAND 196 197
 FT TURN 199 200
 FT HELIX 202 219
 FT TURN 220 220
 FT STRAND 223 226
 FT TURN 227 228
 FT HELIX 229 231
 FT TURN 234 235
 FT HELIX 236 244
 FT STRAND 247 250
 FT HELIX 257 265
 FT TURN 266 266
 FT STRAND 269 271
 FT HELIX 273 283
 FT TURN 286 287
 FT HELIX 290 303
 FT STRAND 307 309
 FT TURN 311 312
 FT TURN 317 318
 FT HELIX 322 325
 FT HELIX 329 341
 FT STRAND 345 349
 FT TURN 350 351
 FT HELIX 352 354
 FT TURN 355 355
 FT TURN 361 364
 FT HELIX 368 371
 FT TURN 372 372

Query Match 100.0%; Score 45; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.24;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8

Db 469 DGNVPVPM 476

RESULT 3
 ID AMYB_ASPAW STANDARD; PRT; 499 AA.
 AC Q02906;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Alpha-amyase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 DE glucanohydrolase B).
 GN AMYB.
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=105351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DKV143E;
 RX MEDLINE=90254827; PubMed=2340591;
 RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
 RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
 RA Berka R.M.;
 RT "Cloning, characterization, and expression of two alpha-amyase genes
 RT from Aspergillus niger var. awamori";
 RL Curr. Genet. 17:203-212(1990).
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: X52756; CAA36967.1; -.
 DR HSSP: P10529; YTA.
 DR InterPro: IPR000461; Alpha_amyase.
 DR Pfam: PF00128; alpha-amyase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 SQ SEQUENCE 499 AA; 54921 MW; 740B96B1BC01A8A CRC64;
 N-LINKED (GLCNAC...) (POTENTIAL).

Query Match 100.0%; Score 45; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.24;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8

Db 469 DGNVPVPM 476

RESULT 4
 AMY_ASPSH

```

ID  AMY_ASPSH  STANDARD;  PRT:  499 AA.
AC  P30292;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY.
OS  Aspergillus shirousami.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
CC  Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
RX  NCBI_TaxID=5070;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92323146; PubMed=1368777;
RA  Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT  "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
RT  expression in Saccharomyces cerevisiae."
RL  Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC  -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -1- COPOLYMER: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D10461; BAA01255.1; -.
DR  PIR: J50663; J50663.
DR  HSSP: P10529; 77RA.
DR  InterPro: IPR000461; Alpha_amylase.
DR  Pfam: PF00128; alpha-amylase; 1.
KM  Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KM  Glycoprotein; Signal.
FT  SIGNAL 1 21
FT  CHAIN 22 499 ALPHA-AMYLASE.
FT  ACT_SITE 227 227
FT  ACT_SITE 231 231
FT  ACT_SITE 251 251 BY SIMILARITY.
FT  ACT_SITE 318 318 BY SIMILARITY.
FT  ACT_SITE 51 59 BY SIMILARITY.
FT  DISULFID 171 185 BY SIMILARITY.
FT  DISULFID 261 304 BY SIMILARITY.
FT  DISULFID 461 496 BY SIMILARITY.
FT  CARBOHD 218 218 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ  SEQUENCE 499 AA; 34652 MW; 1FB7AE50DA01C03F CRC64;

```

```

Query Match      100.0%; Score 45; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY  1 DGNVPPM 8
    |||
Db  469 DGNVPPM 476

```

```

RESULT 5
Y174_HUMAN  STANDARD;  PRT:  364 AA.
AC  P53990;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Hypothetical protein KIAA0174.
GN  KIAA0174.
OS  Homo sapiens (Human).

```

```

OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
PC  TISSUE=Bone marrow;
RX  MEDLINE=96281124; PubMed=8724849;
RA  Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
RT  "Prediction of the coding sequences of unidentified human genes. V.
RT  the coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT  analysis of cDNA clones from human cell line KG-1."
RL  DNA Res. 3:17-24(1996).
CC  -1- SIMILARITY: TO YEAST YNL265C.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D79996; BAA11491.1; -.
DR  Hypothetical protein.
SQ  SEQUENCE 364 AA; 39750 MW; 0DD3C186A52A4380 CRC64;

```

```

Query Match      80.0%; Score 36; DB 1; Length 364;
Best Local Similarity 75.0%; Pred. No. 9.6;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY  1 DGNVPPM 8
    |||
Db  225 DGNVPPM 232

```

```

RESULT 6
ID  GIN2_MYCTU  STANDARD;  PRT:  446 AA.
AC  Q10378;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Probable glutamine synthetase 2 (EC 6.3.1.2) (Glutamate--ammonia
DE  ligase 2).
GN  GIN2 OR RV2222C OR MT2280 OR MTCY427.03C OR MTCY190.33C.
OS  Mycobacterium tuberculosis.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC  Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX  NCBI_TaxID=1773;
RN  [1]
RP  SEQUENCE FROM N.A.
PC  STRAIN=H37RV;
RX  MEDLINE=98295987; PubMed=9634230;
RA  Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA  Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA  Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA  Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA  Horsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA  Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA  Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA  Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT  "Deciphering the biology of Mycobacterium tuberculosis from the
RT  complete genome sequence."
RL  Nature 393:537-544(1998).
RN  [2]
RP  SEQUENCE FROM N.A.
PC  STRAIN=CDC 1551 / Oshkosh;
RX  Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA  Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA  Kolonyak J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA  Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA  Bishai W.;

```

```

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT Laboratory strains.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -1- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC HEXAGON (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z70692; CAA94645.1; -.
DR EMBL: AE007073; AAK46565.1; -.
DR HSSP: P06201; ILGR.
DR TIGR: MT2280; -.
DR TrEMBL: Rv2222c; -.
DR InterPro: IPR001691; GLN_synth.
DR InterPro: IPR001637; GLN_adenyln.
DR Pfam: PF00120; gln-synt; 1.
DR ProDom: PD001057; GLN_adenyln; 1.
DR PROSITE: PS00180; GLN_1; FALSE_NEG.
DR PROSITE: PS00181; GLN_ATP; FALSE_NEG.
DR Ligase: Multigene family; Complete proteome.
SQ SEQUENCE 446 AA; 49507 MW; 86F163FD017829DD CRC64;

Query Match 80.0%; Score 36; DB 1; Length 446;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVP 8
DB 144 DGSVPVPV 151

RESULT 7
XNNZ_CLOTM STANDARD; PRT; 837 AA.
AC P10478;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Endo-1,4-beta-xylanase Z precursor (EC 3.2.1.8) (Xylanase Z)
DE (1,4-beta-D-xylan xylanohydrolase Z).
GN XNNZ.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 10682; PubMed=3139632;
RX MEDLINE=89008072; PubMed=3139632;
RA Grepinet O., Chebrou M.-C., Beguin P.;
RT "Nucleotide sequence and deletion analysis of the xylanase gene
RT (xnnz) of Clostridium thermocellum.";
RN J. Bacteriol. 170:4582-4588(1988).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS) OF 515-837.
RC STRAIN=NCIB 10682;
RX MEDLINE=95393242; PubMed=7664125;
RA Dominguez R., Souchon H., Spitzelli S., Dauter Z., Wilson K.S.,
RA Chavaux S., Beguin P., Alzari P.M.;
RT "A common protein fold and similar active site in two distinct
RT families of beta-glycanases.";
RN Nat. Struct. Biol. 2:569-576(1995).

```

```

CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC linkages in xylans.
CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
CC WELL AS IN OTHER C.THERMOCCELLUM CELLULOSE ENZYMS. THIS DOMAIN
CC MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY F (FAMILY 10 OF GLYCOSYL
CC HYDROLASES).
CC -1- SIMILARITY: CONTAINS 1 XNNZ-TYPE CELLULOSE-BINDING DOMAIN (CBD).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M22624; AAA3286.1; -.
DR PIR: A31842; A31842.
DR PDB: 1XYZ; 29-JAN-96.
DR InterPro: IPR002105; Dockerin_1.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR00379; Est_1ip_thioest_actsite.
DR InterPro: IPR001000; Glyco_hydro_10.
DR Pfam: PF00404; Dockerin_1; 2.
DR Pfam: PF00331; Glyco_hydro_10; 1.
DR PRINTS: PR00134; GLHYDRLASE10.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_2.
DR PROSITE: PS00448; CLOS_CELLULOSE_RPT; 2.
DR PROSITE: PS00591; GLYCOSYL_HYDROL_F10; 1.
DR Xylan degradation; Hydrolase; Glycosidase; Repeat; Signal;
KM 3d-structure.
FT SIGNAL 1 28
FT CHAIN 29 837 POTENTIAL.
FT ACT_SITE 645 645 ENDO-1,4-BETA-XYLANASE Z.
FT ACT_SITE 754 754 PROTON-DONOR.
FT DOMAIN 328 416 NUCLEOPHILE.
FT DOMAIN 430 487 CELLULOSE-BINDING (POTENTIAL).
FT REPEAT 430 453 2 X 24 AA APPROXIMATE REPEATS.
FT REPEAT 464 487 1.
FT REPEAT 488 489 2.
FT DISULFID 783 789 BY SIMILARITY.
SQ SEQUENCE 837 AA; 92262 MW; DD4C29F04D12B6CD CRC64;

Query Match 80.0%; Score 36; DB 1; Length 837;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNVPVP 7
DB 284 DGNTPVP 290

RESULT 8
VCAP_VZVD STANDARD; PRT; 1396 AA.
ID VCAP_VZVD
AC P09245;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major capsid protein (MCP).
DE 40.
GN Varicella-zoster virus (strain Dumas) (VZV).
OS Varicellae; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RN J. Gen. Virol. 67:1759-1816(1986).
RT -1- FUNCTION: MAJOR PROTEIN OF THE ICOSAEDRAL CAPSID.

```

```

CC -! SIMILARITY: BELONGS TO THE HERPESVIRUSES MAJOR CAPSID PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X0A370; CAA27923.1; -
DR PIR: E27341; VCBRA40.
DR InterPro: IPR000912; Herpes_MCP.
DR Pfam: PF03122; Herpes_MCP.1.
DR PRINTS: PR00235; HSVCAPSIDMCP.
DR Coat protein.
DR Coarse protein.
KW SEQUENCE 1396 AA; 154979 MW; E509272D8077242F CRC64;

Query Match
Best Local Similarity 62.5%; Score 36; DB 1; Length 1396;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPPVM 8
Db 595 NCNIPVPL 602

RESULT 9
CHS2_SCHPO STANDARD; PRT; 926 AA.
ID CHS2_SCHPO
AC 074756;
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitin synthase 2 (EC 2.4.1.16) (Chitin-UDP-acetyl-glucosaminyl
DE transferase 2).
DE CHS2 OR SPBC1709.01 OR SPBC1724.17.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M., Rejandream M.A., Barrell B.G., Rieger M.;
RU Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
CC -! CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + ((1,4)-(N-acetyl-
CC beta-D-glucosaminyl)) (N) = UDP + ((1,4)-(N-acetyl-beta-D-
CC glucosaminyl)) (N+1).
CC -! SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -! SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL031856; CAA21311.1; -
DR EMBL: AL031852; CAA21240.1; -
DR InterPro: IPR002923; Chitin_synth.
DR Pfam: PF01644; Chitin_synth.1.
DR ProDom: PD002998; Chitin_synth.1.
KW Transferrase; glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family
SQ SEQUENCE 926 AA; 105578 MW; 7B880751AD53115E CRC64;

```

```

Query Match
Best Local Similarity 75.6%; Score 34; DB 1; Length 926;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPPVM 8
Db 357 DGNVPMOM 364

RESULT 10
SOG_DROME STANDARD; PRT; 1038 AA.
ID SOG_DROME
AC Q24025; Q9VX57;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dorsal-ventral patterning protein Sog (Short gastrulation protein).
GN SOG OR CG9224.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95047398; PubMed=7958919;
RA Francois V., Soliway M., O'Neill J.W., Emery J., Bier E.;
RT "Dorsal-ventral patterning of the Drosophila embryo depends on a
RT putative negative growth factor encoded by the short gastrulation
RT gene."
RL Genes Dev. 8:2602-2616(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt E.G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale U., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadien E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn P.,
RA Durbin K., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.D., Harvey D., Heilmann T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali B.E., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., Meled M.P., Methereson D.,
RA Meruliov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtks R., Tector C.C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."

```

```

RL Science 287:2185-2195(2000).
CC -1- FUNCTION: PUTATIVE NEGATIVE GROWTH FACTOR: ANTAGONIST OF DPP, A
CC PROTEIN INVOLVED IN PATTERNING THE DORSAL REGION AND IN THE
CC DEVELOPMENT OF THE NEUROECTODERM.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: ABOUT THE DORSAL DPP-EXpressING CELLS IN A
CC LATERAL STRIPE 14-16 CELLS WIDE. LATER IN EMBRYOGENESIS IT IS
CC EXPRESSED IN NEUROECTODERM AND IN THE ENDODERM SPACED ALONG THE
CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING GUT.
CC -1- DEVELOPMENTAL STAGE: EMBRYOGENESIS.
CC -1- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.
CC -1- SIMILARITY: CONTAINS 4 VMFC DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; U18774; AAA89117.1; -.
DR EMBL; AE003499; AAF48481.1; -.
DR FlyBase; FBgn0003463; sog.
DR InterPro; IPR001007; VMFC.
DR Pfam; PF00093; vmc; 4.
DR SMART; SM00214; vmc; 3.
DR PROSITE; PS01208; VMFC; 2.
KW Transmembrane; Developmental protein; Repeat; Growth factor;
KW Growth regulation; Signal-anchor.
FT DOMAIN 1 53 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 54 74 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT 1 53 (POTENTIAL).
FT DOMAIN 75 1038 EXTRACELLULAR (POTENTIAL).
FT 100 175 VMFC 1.
FT REPEAT 421 522 SR1.
FT REPEAT 592 668 SR2.
FT REPEAT 677 754 SR3.
FT DOMAIN 742 804 VMFC 2.
FT DOMAIN 830 899 VMFC 3.
FT DOMAIN 939 1020 VMFC 4.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 520 520 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 821 821 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1038 AA; 115514 MW; B0E833AFD79A9037 CRC64;

```

```

Query Match 75.6%; Score 34; DB 1; Length 1038;
Best Local Similarity 85.7%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DGNVFPV 7
Db 723 DNNVFPV 729

```

```

RESULT 11
PCFB_HUMAN STANDARD; PRT; 1654 AA.
AC 094913; 043671;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pre-mRNA cleavage complex II protein Pcf11 (Fragment).
GN Pcf11 OR KIA0824.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE OF 11-1654 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
RN [2]
RP SEQUENCE OF 1-881 FROM N.A.
RA Bouillard J.M., Lacroste F.;
RT "Cloning and sequencing of an Homo sapiens cDNA homologous to the
RT yeast PCF11 gene.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=20514154; PubMed=11060040;
RA de Vries H., Rueggeger U., Huebner W., Friedlein A., Langen H.,
RA Keller W.;
RT "Human pre-mRNA cleavage factor I(m) contains homologs of yeast
RT proteins and bridges two other cleavage factors.";
RL EMBL J. 19:5895-5904(2000).
CC -1- FUNCTION: COMPONENT OF PRE-MRNA CLEAVAGE COMPLEX II.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- CAUTION: REF 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 811.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; AB020631; BAA74847.1; -.
DR EMBL; AF046935; AAC03107.1; ALT_FRAME.
KW mRNA processing; Nuclear protein.
FT NON_TER 1 1
FT DOMAIN 427 535 LYS-RICH.
FT DOMAIN 802 1179 GLY-RICH.
FT DOMAIN 38 41 POLY-ARG.
FT DOMAIN 898 901 POLY-GLY.
FT DOMAIN 941 944 POLY-GLY.
FT DOMAIN 1403 1406 POLY-GLU.
FT DOMAIN 1555 1559 POLY-GLU.
SQ SEQUENCE 1654 AA; 183980 MW; BE29D6E164117BB6 CRC64;

```

```

Query Match 75.6%; Score 34; DB 1; Length 1654;
Best Local Similarity 71.4%; Pred. No. 1,2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 GNVFPVM 8
Db 1303 GNTFAPM 1309

```

```

RESULT 12
RT06_MOUSE STANDARD; PRT; 125 AA.
AC P58064;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial 28S ribosomal protein S6 (MRP-S6).
GN MRPS6 OR RPSM56.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Suzuki T., Terasaki M., Takemoto C., Hanada T., Wada A., Ueda T.,
RA Watanabe K.;
RT "Mammalian mitochondrial ribosome: structural and functional
RT compensation for deficit of RNA with protein components.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE S6P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB049943; BAB40996.1; ALT_INIT.
CC InterPro: IPR000529; RIBOSOMAL_S6.
CC PROSITE: PS01048; RIBOSOMAL_S6: FALSE_NEG.
CC
CC Ribosomal protein; Mitochondrion
CC
CC SEQUENCE 125 AA; 14308 MW; 8140B0BDFDD9025 CRC64;
SQ

```

Query Match
Best Local Similarity 73.3%; Score 33; DB 1; Length 125;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 DGNVPVPM 8
DB 106 DGIIVPVL 113

```

RESULT 13
NIR_PSESP STANDARD; PRT; 379 AA.

```

AC 001537;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE Copper-containing nitrite reductase precursor (EC 1.7.99.3) (Cu-NIR).
GN NIRU
OS Pseudomonas sp. (strain G-179).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93175864; PubMed=8439151;
RA Ye R.W., Fries M.R., Bezborodnikov S.G., Averill B.A., Tiedje J.M.;
RT "Characterization of the structural gene encoding a copper-containing
RT nitrite reductase and Homology of this gene to DNA of other
RT denitrifiers.";
RL Appl. Environ. Microbiol. 59:250-254(1993).
CC -1- CATALYTIC ACTIVITY: 2 nitric oxide + 2 H(2)O + acceptor = 2
CC nitrite + reduced acceptor.
CC -1- COFACTOR: CONTAINS ONE ATOM OF TYPE I COPPER AND ONE ATOM OF TYPE
CC II COPPER. FAD. TYPE I COPPER IS BOUND WITHIN A SINGLE MONOMER,
CC WHILE TYPE 2 COPPER IS HELD BY RESIDUES FROM EACH OF TWO MONOMERS
CC OF THE TRIMER. NITRITE IS BOUND TO THE TYPE 2 COPPER SITE.
CC PSEUDODOXURIN IS THE PHYSIOLOGICAL ELECTRON DONOR FOR THE CU-NIR IN
CC VITRO (BY SIMILARITY).
CC -1- PATHWAY: NITRATE ASSIMILATION (DENITRIFICATION PATHWAY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- DOMAIN: THE TYPE I COPPER SITE IN NIR PLAYS A CRUCIAL ROLE FOR
CC ELECTRON TRANSFER FROM PSEUDODOXURIN TO THE TYPE II COPPER SITE
CC OF NIR, WHICH COMPRISES THE CATALYTIC CENTER OF NIR FOR THE
CC REDUCTION OF NITRITE.
CC -1- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-12 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M97294; AAC79132.1; -.
CC PIR: A48936; A48936.
CC HSSP: P25006; INTF.
CC InterPro: IPR001117; Cu-oxidase.
CC InterPro: IPR001287; Cu-nit rdcase.
CC Pfam: PF00394; Cu-oxidase; 2.
CC PRINTS: PR00695; CUNO2RDFASE.
CC Oxidoreductase; Copper; Metal-binding; Flavoprotein; FAD;
CC Nitrate assimilation; Repeat; Periplasmic; Signal.
CC
CC SIGNAL 1 32
CC CHAIN 33 379 COPPER-CONTAINING NITRITE REDUCTASE.
CC DOMAIN 33 214 PLASTOCYANIN-LIKE 1.
CC DOMAIN 215 379 PLASTOCYANIN-LIKE 2.
CC METAL 134 134 COPPER (TYPE 1) (BY SIMILARITY).
CC METAL 139 139 COPPER (TYPE 2) (BY SIMILARITY).
CC METAL 174 174 COPPER (TYPE 2) (BY SIMILARITY).
CC METAL 175 175 COPPER (TYPE 1) (BY SIMILARITY).
CC METAL 184 184 COPPER (TYPE 1) (BY SIMILARITY).
CC METAL 189 189 COPPER (TYPE 1) (BY SIMILARITY).
CC METAL 345 345 COPPER (TYPE 2) (BY SIMILARITY).
SQ SEQUENCE 379 AA; 40694 MW; 090A3CBF662F62F CRC64;

```

Query Match
Best Local Similarity 73.3%; Score 33; DB 1; Length 379;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 DGNVPVPM 8
DB 104 DGSVPGPM 111

```

RESULT 14
AMVA_ASPNG STANDARD; PRT; 484 AA.

```

AC P56271;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Acid alpha-amylose (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=91002514; PubMed=2207069;
RA Boel E., Brady L., Brzozowski A.M., Derewenda Z., Dodson G.G.,
RA Jensen V.J., Petersen S.B., Swift H., Thim L., Woldike H.F.;
RT "2.1 Å resolution of two enzymes from Aspergillus niger:
RT catalytic binding in alpha-amyloses: an x-ray diffraction study at
RT biochemistry 29:6244-6249(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC PDB: 2AA4; 15-JUL-93.
CC InterPro: IPR000461; Alpha-amyase.
CC Pfam: PF00128; alpha-amyase; 1.
CC Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
CC Glycoprotein; 3D-structure.
CC ACT_SITE 206 206
CC ACT_SITE 210 210
CC ACT_SITE 230 230
CC ACT_SITE 297 297
CC DISULFID 30 38

```



```

FT DISULFID 150 164
FT DISULFID 240 283
FT DISULFID 440 475
FT CARBOHYD 24 24
FT CARBOHYD 157 157
FT CARBOHYD 197 197
SQ SEQUENCE 484 AA; 52935 MW; 04D596E3468056D CRC64;
      N-LINKED (GLCNAC. . .) (POTENTIAL).
      N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

Query Match          73.3%; Score 33; DB 1; Length 484;
Best Local Similarity 71.4%; Pred. No. 51;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 GNVVPM 8
   1:11111
Db 449 GDIPVPM 455

```

```

RESULT 15
VCAP_HSVB STANDARD; PRT; 1376 AA.
ID VCAP_HSVB
AC P28920;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major capsid protein (MCP) (Capsid protein VP5).
GN 42.
OS Equine herpesvirus type 1 (strain Abap) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; Pubmed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RL "The DNA sequence of equine herpesvirus 1.";
CC -1- FUNCTION: MAJOR PROTEIN OF THE ICOSAEDRAL CAPSID.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES MAJOR CAPSID PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M86654; AAB02478.1; -.
DR PIR: H36799; VCBED6.
DR InterPro: IPR000912; Herpes_MCP.
DR Pfam: PF03122; Herpes_MCP.1.
DR PRINTS: PR00235; HSCAPSIDMCP.
KW Coat protein.
SQ SEQUENCE 1376 AA; 152182 MW; C3E866EFAE80AEDB CRC64;

```

```

Query Match          73.3%; Score 33; DB 1; Length 1376;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 DGNVPM 8
   1:1:1:1:
Db 578 NGNIPPL 585

```

Search completed: June 13, 2002, 09:11:57
Job time: 1037 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:53:30 ; Search time 158.68 Seconds
(without alignments)
5.600 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45

Sequence: 1 DGNVPVPM 8

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	100.0	478	15	AAK46065	Mutant alpha-amylase
2	45	100.0	478	16	AAK46065	Mature taka-amylase
3	45	100.0	478	16	AAK46065	Aspergillus oryzae
4	45	100.0	478	16	AAK46065	Aspergillus oryzae
5	45	100.0	478	17	AAK46065	Aspergillus oryzae
6	45	100.0	498	22	AAK46065	Amino acid sequenc
7	36	80.0	314	22	ABG19750	Novel human diago
8	36	80.0	384	22	ABG19751	Novel human diago
9	36	80.0	446	19	AAK37905	Mycobacterium tube
10	36	80.0	619	22	ABK62169	Drosophila melanog
11	36	80.0	837	21	AAK70519	Clostridium thermo

12	35	77.8	64	14	AAK30639	Epitope of human C
13	34	75.6	380	22	ABG19748	Novel human diago
14	34	75.6	585	22	ABK67737	Drosophila melanog
15	34	75.6	759	22	ABG07769	Novel human diago
16	34	75.6	759	22	ABG08475	Novel human diago
17	34	75.6	1038	22	ABK64317	Drosophila melanog
18	33	73.3	57	22	ABK40405	Peptide #7911 enco
19	33	73.3	57	22	ABK40405	Peptide #7911 enco
20	33	73.3	57	22	ABK40405	Peptide #7911 enco
21	33	73.3	57	22	ABK40405	Peptide #7911 enco
22	33	73.3	57	22	ABK40405	Peptide #7911 enco
23	33	73.3	57	22	ABK40405	Peptide #7911 enco
24	33	73.3	57	22	ABK40405	Peptide #7911 enco
25	33	73.3	57	22	ABK40405	Peptide #7911 enco
26	33	73.3	57	22	ABK40405	Peptide #7911 enco
27	33	73.3	57	22	ABK40405	Peptide #7911 enco
28	33	73.3	57	22	ABK40405	Peptide #7911 enco
29	33	73.3	57	22	ABK40405	Peptide #7911 enco
30	33	73.3	57	22	ABK40405	Peptide #7911 enco
31	33	73.3	57	22	ABK40405	Peptide #7911 enco
32	33	73.3	57	22	ABK40405	Peptide #7911 enco
33	33	73.3	57	22	ABK40405	Peptide #7911 enco
34	33	73.3	57	22	ABK40405	Peptide #7911 enco
35	33	73.3	57	22	ABK40405	Peptide #7911 enco
36	33	73.3	57	22	ABK40405	Peptide #7911 enco
37	33	73.3	57	22	ABK40405	Peptide #7911 enco
38	33	73.3	57	22	ABK40405	Peptide #7911 enco
39	33	73.3	57	22	ABK40405	Peptide #7911 enco
40	33	73.3	57	22	ABK40405	Peptide #7911 enco
41	33	73.3	57	22	ABK40405	Peptide #7911 enco
42	33	73.3	57	22	ABK40405	Peptide #7911 enco
43	33	73.3	57	22	ABK40405	Peptide #7911 enco
44	33	73.3	57	22	ABK40405	Peptide #7911 enco
45	33	73.3	57	22	ABK40405	Peptide #7911 enco

ALIGNMENTS

RESULT 1	
AAK46065	standard; protein: 478 AA.
AC	AAK46065;
XX	
XX	18-JUL-1994 (first entry)
DE	Mutant alpha-amylase.
XX	
KW	Methionine substitution; stability; activity; detergent;
KW	dishwashing agents; liquifaction agents.
XX	
OS	Aspergillus oryzae.
XX	
PN	WO9402597-A.
PD	03-FEB-1994.
XX	
PF	06-JUL-1993; 93WO-DK00230.
XX	
PR	23-JUL-1992; 92DK-0000946.
PR	16-DEC-1992; 92DK-0001503.
PR	15-MAR-1993; 93DK-0000292.
XX	
PA	(NOVO) NOVO-NORDISK AS.
XX	
PI	Bisgard-frantzen H, Svendsen A;
XX	
DR	WPI, 1994-048855/06.
XX	
PT	Mutant alpha-amylase from Bacillus species comprising a
PT	methionine substitution - with improved stability and activity at
PT	low pH, for use in detergents, dishwashing agents and

PT liquefaction agents
XX
PS Claim 1; Page 7; 20pp; English.
XX
CC The sequence os that of the Aspergillus oryzae alpha amylase, sold
CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC be mutated by substitution of one or more of its methionine residues
CC for any amino acid other than cysteine. The mutant alpha-amylase
CC exhibits a better activity level and better stability in the
CC presence of oxidizing agents than previous mutant alpha amylases,
CC and improved thermostability at moderately low pH. The enzyme can
CC be used as an additive for detergents, dishwashing agents and
CC liquefaction agents.
XX
SQ Sequence 478 AA;
XX
Query Match 100.0%; Score 45; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DGNVPVPM 8
DB 448 dgnvpvpm 455
XX
RESULT 2
ID AAR79025 standard; protein: 478 AA.
XX
AC AAR79025;
XX
DT 22-MAR-1996 (first entry)
XX
DE Mature taka-amylase A.
XX
KW Wild type; neopullulanase; B. stearothermophilus; mutant;
KW Food industry; modification; hydrophobicity; replacement
KW insertion; deletion.
XX
OS Aspergillus oryzae.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
PN JP07177891-A.
XX
PD 18-JUL-1995.
XX
PF 31-OCT-1994; 94JP-0288658.
XX
PR 12-NOV-1993; 93JP-0306096.
XX
PA (EZAKI) EZAKI GLICO CO.
XX
PA (NIDE) NEC CORP.
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site - increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
PS Disclosure; Page 10-11; 18pp; English.
XX
CC This sequence represents the mature form of taka-amylase from A. oryzae.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the

CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC hydrophobic amino acid from the selected site. The method was used
CC to produce neopullulanases Y377F, S422V and M375L.
XX
SQ Sequence 478 AA;
XX
Query Match 100.0%; Score 45; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DGNVPVPM 8
DB 447 dgnvpvpm 454
XX
RESULT 3
ID AAR72450 standard; protein: 478 AA.
XX
AC AAR72450;
XX
DT 01-DEC-1995 (first entry)
XX
DE Aspergillus oryzae alpha amylase (mature protein).
XX
KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
KW Bacillus stearothermophilus; dyeing; bleaching; scouring; textile;
KW thermostable.
XX
OS Aspergillus oryzae.
XX
PN WO9510603-A.
XX
PD 20-APR-1995.
XX
PF 05-OCT-1994; 94WO-DK00370.
XX
PR 08-OCT-1993; 93DK-0001133.
PR 02-FEB-1994; 94DK-0000140.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Svendsen A, Thellersen M;
PI Van der zee P, Bisgardfrantzen H, Borchert T;
XX
DR WPI; 1995-161790/21.
XX
PT New Bacillus derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance
XX
PS Disclosure; Page 75-76; 105pp; English.
XX
CC Variant alpha amylase enzymes which have improved washing and/or
CC as detergent additives. The enzymes have one or more amino acid
CC residues added, deleted or substituted. The variants can also be
CC used for textile desizing prior to scouring, bleaching and dyeing.
CC The variants have improved thermostability, acid/alkaline stability;
CC low temperature optimum; pH optimum; higher hydrolysis velocity and
CC improved tolerance to other composition constituents, e.g. oxidation
CC agents.
XX
SQ Sequence 478 AA;
XX
Query Match 100.0%; Score 45; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DGNVPVPM 8
DB 447 dgnvpvpm 454

Dd		448 dgnvpypm 455	
RESULT	1		
ID	AAR78270	standard; Protein; 478 AA.	
XX	AAR78270;		
AC			
XX			
DT	17-JAN-1996	(first entry)	
DE	Aspergillus oryzae alpha amylase (mature protein).		
XX			
KW	Alpha amylase; oxidation; desizing; bleaching; scouring; fabric starch; thermostable; methionine; Bacillus licheniformis; Bacillus amyloliquefaciens; Bacillus stearothermophilus; Aspergillus oryzae.		
XX			
OS	Aspergillus oryzae.		
PX			
PN	WO9521247-A1.		
XX			
PD	10-AUG-1995.		
XX			
PF	05-OCT-1994; 94WO-DK00371.		
XX			
PR	02-FEB-1994; 94DK-0000141.		
XX			
PA	(NOVO) NOVO-NORDISK AS.		
PI	Marcher D, Nilsson TE, Pedersen HH, Toft AH;		
DR	WPI; 1995-283767/37.		
XX			
PT	Use of an oxidation stable alpha-amylase - for simultaneous desizing and bleaching or scouring of fabrics contg. starch or starch derivs.		
PS	Disclosure; Page 25-26; 37pp; English.		
XX			
CC	Oxidation stable alpha amylases can be used for the simultaneous desizing and bleaching or scouring of a fabric comprising starch or starch derivatives. They exhibit a better heat stability, especially in the presence of oxidising agents. They are obtained from a parent alpha amylase by replacing one or more methionine residues with any amino acid different from Cys or Met, preferably Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is pref. derived from a Bacillus species, although alpha amylases of fungal origin can also be used. This sequence is the wild type (unmodified) alpha amylase of Aspergillus oryzae.		
CC			
CC			
XX			
SQ	Sequence 478 AA;		
Query Match	100.0%; Score 45; DB 16; Length 478;		
Best Local Similarity	100.0%; Pred. No. 3.2;		
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0.		
QY	1 DGNVPPM 8 		
Dd	448 dgnvpypm 455		
RESULT	5		
AAWI4500			
ID	AAWI4500 standard; protein; 478 AA.		
XX			
AC	AAWI4500;		
XX			
DT	04-JUN-1997 (first entry)		
XX			
DE	Aspergillus oryzae alpha-amylase (mature protein).		
XX			
KW	alpha-amylase; mature; Bacillus licheniformis; terramyl; fungamyl;		

KW	Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;
KW	calcium dependency; substrate binding; stability; pH optimum;
KM	thermostability; cleavage; oligosaccharide substrate; dishwashing;
KW	washing; detergent additive; fabric desizing; starch liquefaction;
KM	sweetener; ethanol production; variant.
XX	
OS	Aspergillus oryzae.
XX	
FH	Key
FH	location/qualifiers
FT	Misc-difference 13..45
FT	/label= loop 1 modification region
FT	/note= "at least one amino acid residue of a parent
FT	alpha-amylase (used as a template for a variant)
FT	corresponding to 7-23 of AAM14499 is deleted or
FT	replaced with a fragment corresponding to
FT	this fragment; claim 33"
FT	
FT	Misc-difference 14..40
FT	/label= loop 1 modification region
FT	/note= "preferred region where at least one amino acid
FT	residue of a parent alpha-amylase (used as a
FT	template for a variant) corresponding to 8-18
FT	of AAM14499 is deleted or replaced with a
FT	fragment corresponding to this fragment; claim
FT	35"
FT	35" Misc-difference 28..42
FT	/label= loop 1 modification region
FT	/note= "at least one amino acid residue of a parent
FT	alpha-amylase (used as a template for a variant),
FT	corresponding to 12-19 of AAM14499 is deleted or
FT	replaced with a fragment corresponding to
FT	this fragment; claim 30"
FT	
FT	Misc-difference 32..38
FT	/label= loop 1 modification region
FT	/note= "preferred region where at least one amino acid
FT	residue of a parent alpha-amylase (used as a
FT	template for a variant) corresponding to 14-15
FT	of AAM14499 is deleted or replaced with a
FT	fragment corresponding to this fragment; claim
FT	32"
FT	32" Misc-difference 66..84
FT	/label= loop 2 modification region
FT	/note= "at least one amino acid residue of a parent
FT	alpha-amylase (used as a template for a variant)
FT	corresponding to 44-57 of AAM14499 is deleted or
FT	replaced with a fragment corresponding to
FT	this fragment; claim 18"
FT	
FT	Misc-difference 70..78
FT	/label= loop 2 modification region
FT	/note= "preferred region where at least one amino acid
FT	residue of a parent alpha-amylase (used as a
FT	template for a variant) corresponding to 48-51
FT	of AAM14499 is deleted or replaced with a
FT	fragment corresponding to this fragment; claim
FT	20"
FT	20" Misc-difference 98..110
FT	/label= loop 3 modification region
FT	/note= "at least one amino acid residue of a parent
FT	alpha-amylase (used as a template for a variant),
FT	corresponding to 117-185 of AAM14499 is deleted
FT	or replaced with a fragment corresponding to
FT	this fragment; claim 24"
FT	
FT	Misc-difference 102..206
FT	/label= loop 3 modification region
FT	/note= "preferred region where at least one amino acid
FT	residue of a parent alpha-amylase (used as a
FT	template for a variant) corresponding to 121-181 of
FT	AAM14499 is deleted or replaced with a
FT	fragment corresponding to this fragment; claim
FT	26"
FT	26" Misc-difference 121..181
FT	/note= "at least one amino acid residue of a parent
FT	alpha-amylase (used as a template for a variant),
FT	corresponding to this fragment is deleted or
FT	replaced with a fragment corresponding to
FT	102-206 of AAM14499; claim 41"
FT	
FT	Misc-difference 121..174
FT	/note= "preferred region where at least one amino acid

RESULT 7
ABG19750
ID ABG19750 standard; Protein; 314 AA.
XX
AC ABG19750;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19741.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS83937.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 50109; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 314 AA;

Query Match 80.0%; Score 36; DB 22; Length 314;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 DGNVPVPM 8
1111111
Db 112 dgtvpm 119

RESULT 8
ABG19751
ID ABG19751 standard; Protein; 384 AA.
XX
AC ABG19751;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19742.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS83938.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 50110; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 384 AA;

Query Match 80.0%; Score 36; DB 22; Length 384;
Best Local Similarity 75.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 DGNVPVPM 8
1111111
Db 233 dgtvpm 240

RESULT 9

AAW37905
ID AAW37905 standard; Protein: 446 AA.
XX
XX AAW37905;
AC
XX
XX 11-SEP-1998 (first entry)
DT
XX
XX Mycobacterium tuberculosis glutamine synthetase II.
DE
XX
XX N-terminal; extracellular protein; abundant; tuberculosis; Tb;
KM pathogen; disease; treatment; mycobacterial; drug resistant;
KW glutamine synthetase II.
XX
XX Mycobacterium tuberculosis.
OS
XX
XX W09814613-A1.
PN
XX
XX 09-APR-1998.
PD
XX
XX 29-SEP-1997; 97WO-US17756.
PF
XX
XX 02-OCT-1996; 96US-0724814.
PR
XX
XX (REGC) UNIV CALIFORNIA.
PA
XX
XX Harth G, Horwitz MA;
PI
XX
XX WPI: 1998-240107/21.
DR
XX
XX N-PSDB: AAV29071.
PT
XX
XX Treatment of infections caused by pathogenic organisms - by
PT interfering with the functional activity of an extracellular enzyme
PT identified as necessary for the growth or survival of the organism
XX
XX Example 16; Page 31-32; 88pp; English.
PS
XX
XX The sequence is that of glutamine synthetase II which was used as
CC part of a method for the treatment of mammalian disease conditions
CC associated with infection by a pathogenic organism (PO) comprising:
CC (a) identifying at least one extracellular enzyme necessary for the
CC growth or survival of the PO; and (b) interfering with the functional
CC activity of the extracellular enzyme to inhibit the growth or survival
CC of the PO. The method can be used to treat diseases caused by POs such
CC as bacteria, protozoa, viruses and fungi. In particular they can
CC be used for the treatment of diseases caused by intracellular
CC mycobacterial pathogens including M. tuberculosis, M. bovis,
CC M. avium, M. kansasii, M. fortuitum, M. chelonae, M. marinum,
CC M. scrofulaceum, M. leprae, M. africanum, M. ulcerans and
CC M. microti. The methods may be particularly useful for treating
CC drug resistant mycobacterium. Using the methods it is possible
CC to treat disease conditions associated with sequestered
CC intracellular PO without harming the infected host.
XX
XX
SQ Sequence 446 AA;

Query Match 80.0%; Score 36; DB 19; Length 446;
Best Local Similarity 75.0%; Pred. NO. 1.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 144 dgsyppvpv 151

RESULT 10
ID AAB62169 standard; Protein: 619 AA.
XX
XX AAB62169;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX

DE Drosophila melanogaster polypeptide SEQ ID NO 13299.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX W0200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI: 2001-656860/75.
DR
XX
XX N-PSDB: ABL06272.
PT
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
XX Disclosure; SEQ ID NO 13299; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB157737-AB172072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 619 AA;

Query Match 80.0%; Score 36; DB 22; Length 619;
Best Local Similarity 75.0%; Pred. NO. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 207 dqxvplpm 214

RESULT 11
ID AAY70519 standard; Protein: 837 AA.
XX
XX AAY70519;
AC
XX
XX 04-JUL-2000 (first entry)
DT
XX
XX Clostridium thermocellum xylanase Z.
DE
XX
XX xylanase Z; xynZ; cellulosome; feruloyl esterase; phenolic acid esterase;
KM thermostable; ferulic acid; wheat bran; agricultural byproduct; treat;
KW grass; paper and pulp industry; feed processing; food additive;
KW plant cell wall material; degradation.
XX
XX Clostridium thermocellum.
OS
XX
XX
XX Key Location/Qualifiers
FH 287..296
FT Region /label= Pro-/Thr_-region


```

FT Binding-site 289..400
FT /label= Family_VI_cellulose_binding_domain
FT Region 430..453
FT /note= "conserved region"
FT Region 464..487
FT /note= "conserved region"
FT Domain 20..286
FT /note= "This region is a part of multi-domain structure
FT of XynZ. It is specifically claimed and
FT contains feruloyl esterase"
FT Domain 20..307
FT /note= "This region is a part of multi-domain structure
FT of XynZ. It is specifically claimed and
FT contains feruloyl esterase"
FT Domain 20..421
FT /note= "This region is a part of multi-domain structure
FT of XynZ. It is specifically claimed and
FT contains feruloyl esterase"
FT Domain 69..286
FT /note= "specifically exemplified feruloyl esterase"
FT Domain 69..307
FT /note= "specifically exemplified feruloyl esterase"
FT Domain 69..421
FT /note= "specifically exemplified feruloyl esterase"
FT Domain 69..421
FT /note= "specifically exemplified feruloyl esterase"
XX WO200014243-A1.
XX 16-MAR-2000.
XX 03-SEP-1999; 99WO-US20304.
XX 04-SEP-1998; 98US-0099136.
XX (UYGE-) UNIV GEORGIA RES FOUND INC.
XX Blum DL, Kataeva I, Li X, Ljungdahl LG;
XX WPI; 2000-256991/22.
XX N-PSDB; AAZ51818.
XX New recombinant DNA molecule comprising a sequence encoding feruloyl
XX esterase protein, useful for treating grasses and other plant materials
XX used in pulp and paper industries, feed processing and food additives -
XX Claim 23; Pages 37-38; 105pp; English.
XX The present sequence is a xylanase Z (XynZ) from
XX Clostridium thermocellum. XynZ is an enzymatic component of
XX C. thermocellum cellulosome and has a multi-domain structure which
XX includes a dockerin domain, a catalytic xylanase domain,
XX a family VI cellulose binding domain and a domain of
XX unknown function. The unknown domain in the N-terminal region of XynZ
XX has been found to contain feruloyl (phenolic acid) esterase which
XX is involved in the degradation of plant cell wall material. The novel
XX feruloyl esterase is thermostable, easy to purify, has high temperature
XX optima and stable over a wide pH range. The enzyme is used for
XX producing ferulic acid from wheat bran or agricultural byproducts,
XX treating grasses or other plant materials used in the pulp and paper
XX industries, in feed processing and as a food additive.
XX Sequence 837 AA;
SQ

```

```

Query Match 80.0%; Score 36; DB 21; Length 837;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 DGNVPVP 7
   ||| |||
DB 284 dgnvpvp 290

```

RESULT 12

```

AAR30639
ID AAR30639 standard; peptide; 64 AA.
XX
AC AAR30639;
XX
DT 06-MAY-1993 (first entry)
XX
DE Epitope of human CENP-B.
XX
KM Centromere protein; antibodies; polypeptide.
XX
OS Homo sapiens.
XX
PN JP04334398-A.
XX
PD 20-NOV-1992.
XX
PF 08-MAY-1991; 91JP-0102517.
XX
PR 08-MAY-1991; 91JP-0102517.
XX
PA (DAIK ) DAIKIN KOGYO KK.
XX
DR WPI; 1993-005542/01.
XX
PT Human centromere antigen polypeptide - for detection of human
PT antibodies and identification of disease
XX
XX Claim 1; Page 2; 15pp; Japanese.
XX
XX The polypeptide constitutes an epitope of the human centromere protein
XX B (CENP-B) and can be produced by standard recombinant DNA techniques.
XX The polypeptide can be used to detect anti-human centromere antibodies.
XX The type of diseases of a patient having the antibody can be exactly
XX classified using the polypeptide.
XX See also AAR30640-I.
XX
XX Sequence 64 AA;
SQ

```

```

Query Match 77.8%; Score 35; DB 14; Length 64;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 DGNVPVP 7
   ||| |||
DB 1 dgdvpvp 7

```

```

RESULT 13
ID ABG19748 standard; Protein; 380 AA.
XX
AC ABG19748;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19739.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX

```

PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS83935.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20: SEQ ID No 50107; 103pp: English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 380 AA;
 XX
 Query Match 75.6%; Score 34; DB 22; Length 380;
 Best Local Similarity 75.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DGNVPVPM 8
 II I I I I I
 Db 293 dgmvpmpm 300
 XX
 RESULT 14
 ABG67737
 ID ABG67737 standard; Protein: 585 AA.
 XX
 AC ABG67737;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 30003.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KM
 OS Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 PN
 XX
 PD 27-SEP-2001.
 PD
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 PF
 PR 23-MAR-2000; 2000US-191637P.
 PR
 PR 11-JUL-2000; 2000US-0614150.
 PR
 XX
 PS (PEKE) PE CORP NY.

XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB: ABL11840.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID No 30003; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABG57737-ABG72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 585 AA;
 XX
 Query Match 75.6%; Score 34; DB 22; Length 585;
 Best Local Similarity 71.4%; Pred. No. 4.1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GNVVPVPM 8
 II I I I I I
 Db 144 gnapipm 150
 XX
 RESULT 15
 ABG07769
 ID ABG07769 standard; Protein: 759 AA.
 XX
 AC ABG07769;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #7760.
 XX
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW
 OS Homo sapiens.
 OS
 PN WO200175067-A2.
 PN
 XX
 PD 11-OCT-2001.
 PD
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 PF
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR
 PR 23-AUG-2000; 2000US-0649167.
 PR
 XX
 PA (HYSE-) HYSEQ INC.
 PA
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS71956.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20: SEQ ID No 38128; 103pp: English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. Abg00010-Abg30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 759 AA;

Query Match 75.6%; Score 34; DB 22; Length 759;
Best Local Similarity 71.4%; Pred. No. 5.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 7
||:|:|:|
Db 402 dgsypip 408

Search completed: June 13, 2002, 08:53:31
Job time: 281 sec

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:32 ; Search time 57.95 Seconds
(without alignments)
3.372 Million cell updates/sec

Title: US-09-710-339-2_COPY_448_455

Perfect score: 41

Sequence: 1 YTAGQQLT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/iaa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	41	100.0	478	1	US-08-720-899-7
2	41	100.0	478	1	US-08-459-610-7
3	41	100.0	478	2	US-08-343-804-7
4	41	100.0	478	2	US-08-339-715A-2
5	41	100.0	478	2	US-08-600-908A-10
6	41	100.0	478	3	US-08-683-838A-10
7	41	100.0	478	4	US-09-182-859-7
8	31	75.6	342	1	US-08-499-215-2
9	30	73.2	109	2	US-08-585-585A-1
10	30	73.2	109	2	US-08-249-037C-1
11	30	73.2	109	2	US-08-788-622B-1
12	30	73.2	109	3	US-08-788-622B-1
13	30	73.2	631	3	US-08-814-052-8
14	30	73.2	631	3	US-08-812-829-8
15	29	70.7	177	3	US-08-824-692-30
16	29	70.7	229	3	US-08-824-692-31
17	29	70.7	290	3	US-08-824-692-29
18	29	70.7	505	1	US-08-220-603A-10
19	29	70.7	1463	1	US-08-220-603A-11
20	28	68.3	441	1	US-08-403-866-10
21	27	65.9	214	3	US-08-960-507-21
22	27	65.9	215	4	US-09-247-373B-40
23	27	65.9	263	1	US-08-407-544-2
24	27	65.9	264	1	US-08-719-697-8
25	27	65.9	264	4	US-08-727-616A-8
26	27	65.9	265	2	US-08-719-697-2
27	27	65.9	265	2	US-08-719-697-2
27	27	65.9	265	2	US-08-719-697-2

28	27	65.9	265	4	US-08-727-616A-2	Sequence 2, Appli
29	27	65.9	265	4	US-08-727-616A-6	Sequence 6, Appli
30	27	65.9	286	1	US-07-721-775A-2	Sequence 2, Appli
31	27	65.9	286	1	US-08-339-658-2	Sequence 2, Appli
32	27	65.9	286	1	US-08-346-333-16	Sequence 16, Appli
33	27	65.9	286	2	US-08-719-697-4	Sequence 4, Appli
34	27	65.9	286	4	US-09-263-933-7	Sequence 7, Appli
35	27	65.9	286	4	US-09-263-933-14	Sequence 14, Appli
36	27	65.9	286	4	US-08-727-616A-4	Sequence 21, Appli
37	27	65.9	286	4	US-09-025-769B-265	Sequence 265, App
38	27	65.9	286	4	US-09-025-769B-362	Sequence 362, App
39	27	65.9	286	5	PCT-US91-07506-16	Sequence 16, Appli
40	27	65.9	286	4	US-09-025-769B-285	Sequence 285, App
41	27	65.9	299	4	US-09-025-769B-298	Sequence 298, App
42	27	65.9	299	4	US-09-025-769B-300	Sequence 300, App
43	27	65.9	406	4	US-09-108-020-4	Sequence 4, Appli
44	27	65.9	406	4	US-09-108-020-4	Sequence 4, Appli
45	27	65.9	406	4	US-09-108-020-4	Sequence 42, Appli

ALIGNMENTS

RESULT 1
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchardt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720, 899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343, 804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowrey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054, 214-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7
Query Match 100.0%; Score 41; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
| | | | |
Db 428 YTAGOOLT 435

```

RESULT 2
US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459, 610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343, 804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-459-610-7

Query Match 100.0%; Score 41; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
| | | | |
Db 428 YTAGOOLT 435

RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5850837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
```

```

; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58308370 No. 5830837disk of No. 5830837th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343, 804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-343-804-7

Query Match 100.0%; Score 41; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
| | | | |
Db 428 YTAGOOLT 435

RESULT 4
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michiyo
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339, 715A
; FILING DATE: 14-NOV-1994
```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Meiser, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 41; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
DB 427 YTAGOOLT 434

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ. ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-600-908A-10

Query Match 100.0%; Score 41; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
DB 428 YTAGOOLT 435

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ. ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-683-838A-10

Query Match 100.0%; Score 41; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
DB 428 YTAGOOLT 435

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgard-Frantzen, Henrik

;; TITLE OF INVENTION: Alpha-Amylase Mutants
;; FILE REFERENCE: 4796.204-US
;; CURRENT APPLICATION NUMBER: US/09/182,859
;; CURRENT FILING DATE: 1998-10-29
;; EARLIER APPLICATION NUMBER: 0515/96
;; EARLIER FILING DATE: 1996-04-30
;; EARLIER APPLICATION NUMBER: 0712/96
;; EARLIER FILING DATE: 1996-06-28
;; EARLIER APPLICATION NUMBER: 0775/96
;; EARLIER FILING DATE: 1996-07-11
;; EARLIER APPLICATION NUMBER: 1263/96
;; EARLIER FILING DATE: 1996-11-08
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 7
;; LENGTH: 478
;; TYPE: PRT
;; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 41; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGQOLT 8
Db 428 YTAGQOLT 435

RESULT 8
US-08-499-215-2
; Sequence 2, Application US/08499215
; Patent No. 5612204
; GENERAL INFORMATION:
; APPLICANT: Saeki, Hisashi
; APPLICANT: Miura, Akira
; TITLE OF INVENTION: BIOLOGICAL DEGRADATIVE
; TITLE OF INVENTION: TREATMENT OF CHLORINE-SUBSTITUTED ETHYLENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Miron, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue
; CITY: N.W.
; STATE: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/499,215
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP Hei-6-179689
; FILING DATE: 08-JUL-1994
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 342 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-499-215-2

Query Match 75.6%; Score 31; DB 1; Length 342;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGQOLT 8
Db 34 YTAGQOLT 41

RESULT 9
US-08-585-585A-1
; Sequence 1, Application US/08585585A
; Patent No. 5874308
; GENERAL INFORMATION:
; APPLICANT: Kilburn, Douglas G.
; APPLICANT: Humphries, Keith R.
; APPLICANT: Doherty, James G.
; APPLICANT: Jarvis, Eric
; APPLICANT: Alimonti, Judie
; TITLE OF INVENTION: Compositions and methods for modulating
; TITLE OF INVENTION: cell proliferation using growth factor-polysaccharide
; TITLE OF INVENTION: binding fusion proteins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rae-Venter Law Group
; STREET: Box 60039
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/585,585A
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kung, Viola T.
; REGISTRATION NUMBER: 41,131
; REFERENCE/DOCKET NUMBER: CBPT.016.0005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)328-4400
; TELEFAX: (650)328-4477
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-585-585A-1

Query Match 73.2%; Score 30; DB 2; Length 109;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGQOLT 7
Db 40 YTAGQOLT 46

RESULT 10
US-08-249-037C-1
; Sequence 1, Application US/08249037C
; Patent No. 5928917
; GENERAL INFORMATION:
; APPLICANT: Kilburn, Douglas G.
; APPLICANT: Miller, Robert C.
; APPLICANT: Warren, Richard A.J.
; APPLICANT: Gilkes, Neil R.
; TITLE OF INVENTION: Polysaccharide binding fusion proteins
; TITLE OF INVENTION: and conjugates
; NUMBER OF SEQUENCES: 21

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Rae-Venter Law Group, P.C.
;; STREET: P.O.Box 60039
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: U.S.
;; ZIP: 94306
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/249,037C
;; FILING DATE: 24-MAY-1994
;; CLASSIFICATION: 435
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/865,095
;; FILING DATE: 08-APR-1992
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/603,987
;; FILING DATE: 25-OCT-1990
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/216,794
;; FILING DATE: 08-JUL-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kung, Viola T.
;; REGISTRATION NUMBER: 41,131
;; REFERENCE/DOCKET NUMBER: CBDT.002.04US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (650)328-4400
;; TELEFAX: (650)328-4477
;;
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 109 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
;; US-08-249-037C-1

Query Match 73.2%; Score 30; DB 2; Length 109;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOL 7
DB 40 YTAGORI 46

RESULT 11
US-08-788-622B-1
; Sequence 1, Application US/08788622B
; Patent No. 5962289
; GENERAL INFORMATION:
; APPLICANT: Kilburn, Douglas G.
; APPLICANT: Miller, Robert C.
; APPLICANT: Warren, Richard A.J.
; APPLICANT: Gilkes, Neil R.
; TITLE OF INVENTION: Polysaccharide binding fusion proteins
; TITLE OF INVENTION: and conjugates
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rae-Venter Law Group, P.C.
; STREET: P.O.Box 60039
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/788,622B
;; FILING DATE: January 23, 1997
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/249,037
;; FILING DATE: 24-MAY-1994
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/865,095
;; FILING DATE: 08-APR-1992
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/603,987
;; FILING DATE: 25-OCT-1990
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/216,794
;; FILING DATE: 08-JUL-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kung, Viola T.
;; REGISTRATION NUMBER: 41,131
;; REFERENCE/DOCKET NUMBER: CBDT.002.06US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (650)328-4400
;; TELEFAX: (650)328-4477
;;
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 109 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
;; US-08-788-622B-1

Query Match 73.2%; Score 30; DB 2; Length 109;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOL 7
DB 40 YTAGORI 46

RESULT 12
US-08-788-621B-1
; Sequence 1, Application US/08788621B
; Patent No. 6124117
; GENERAL INFORMATION:
; APPLICANT: Kilburn, Douglas G.
; APPLICANT: Miller, Robert C.
; APPLICANT: Warren, Richard A.J.
; APPLICANT: Gilkes, Neil R.
; TITLE OF INVENTION: Polysaccharide binding fusion proteins
; TITLE OF INVENTION: and conjugates
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rae-Venter Law Group, P.C.
; STREET: P.O.Box 60039
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,621B
; FILING DATE: January 23, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/249,037
; FILING DATE: 24-MAY-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/865,095
FILING DATE: 08-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,987
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/216,794
FILING DATE: 08-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Kung, Viola T.
REGISTRATION NUMBER: 41,131
REFERENCE/DOCKET NUMBER: CDDT.002.0505
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)328-4400
TELEFAX: (650)328-4477
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-788-621B-1

Query Match 73.2%; Score 30; DB 3; Length 109;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOL 7
|||||:
DB 40 YTAGORI 46

RESULT 13
US-08-814-052-8
Sequence 8, Application US/08814052
Patent No. 6015783
GENERAL INFORMATION:
APPLICANT: von der Osten, Claus
APPLICANT: Cherry, Joel R.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Vind, Jesper
APPLICANT: Rasmussen, Michael Dolberg
TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING
TITLE OF INVENTION: OR STRAINS FROM CELLULOSIC FABRIC
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 60157830 No. 6015783disk of No. 6015783th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/814,052
FILING DATE: 06-MAR-1997
CLASSIFICATION: 510
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4684.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6015783e
US-08-814-052-8

Query Match 73.2%; Score 30; DB 3; Length 631;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOL 7
|||||:
DB 73 YTAGORI 79

RESULT 14
US-08-812-829-8
Sequence 8, Application US/08812829
Patent No. 6017751
GENERAL INFORMATION:
APPLICANT: von der Osten, Claus
APPLICANT: Bjornvad, Mads E.
APPLICANT: Vind, Jesper
APPLICANT: Rasmussen, Michael Dolberg
TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING
TITLE OF INVENTION: OR STRAINS FROM CELLULOSIC FABRIC
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 60177510 No. 6017751disk of No. 6017751th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,829
FILING DATE: 06-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4690.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6017751e
US-08-812-829-8

Query Match 73.2%; Score 30; DB 3; Length 631;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOL 7
|||||:
DB 73 YTAGORI 79

```

RESULT 15
US-08-824-692-30
; Sequence 30, Application US/08824692
; Patent No. 6017703
; GENERAL INFORMATION:
; APPLICANT: Kinders, Robert J.
; APPLICANT: Enfield, David L.
; APPLICANT: Hass, G. Michael
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
; TITLE OF INVENTION: FOR OR MODULATING A TUMOR ASSOCIATED ANTIGEN
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,692
; FILING DATE: 08-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 130001.404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 622-6031
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-824-692-30

```

```

Query Match 70.7%; Score 29; DB 3; Length 177;
Best Local Similarity 62.5%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 YTAGOOLT 8
| | | | |
Db 63 YKAGEQVT 70

```

Search completed: June 13, 2002, 08:54:33
Job time: 318 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:15 ; Search time 125.81 Seconds
(without alignments)
11.000 Million cell updates/sec

Title: US-09-710-339-2_COPY_448_455

Perfect score: 41
Sequence: 1 YTAGQQLT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	499	3 Q96TH4	Q96TH4 aspergillus
2	34	82.9	849	16 Q9KND8	Q9KND8 vibrio chol
3	33	80.5	420	16 Q982J7	Q982J7 rhizobium 1
4	32	78.0	192	12 Q9YWK5	Q9YWK5 buzura supp
5	32	78.0	211	5 Q9GPI6	Q9GPI6 anopheles g
6	32	78.0	333	12 Q99H12	Q99H12 heliocoverp
7	32	78.0	390	12 Q9YWS6	Q9YWS6 lymantria d
8	32	78.0	394	12 Q9PYW5	Q9PYW5 xestia c-ni
9	32	78.0	399	12 Q9IML8	Q9IML8 choristoneu
10	32	78.0	399	12 Q92407	Q92407 bombyx mori
11	32	78.0	400	12 Q91G14	Q91G14 epiphyas po
12	32	78.0	400	12 Q91822	Q91822 spodoptera
13	32	78.0	412	12 Q91BX7	Q91BX7 heliocoverp
14	32	78.0	422	12 Q91B15	Q91B15 cydia pomon
15	32	78.0	460	12 Q91EX7	Q91EX7 cydia pomon
16	32	78.0	541	5 Q93W09	Q93W09 glardia lam

17	32	78.0	764	4 Q9P2H3	Q9P2H3 homo sapien
18	32	78.0	1250	5 Q23409	Q23409 caenorhabdi
19	32	78.0	2119	13 Q90X47	Q90X47 brachydanio
20	31	75.6	214	2 Q91OE0	Q91OE0 streptomyc
21	31	75.6	221	16 Q9JTW7	Q9JTW7 neisseria m
22	31	75.6	223	5 Q9V8H2	Q9V8H2 drosophila
23	31	75.6	242	9 Q03913	Q03913 bacterioph
24	31	75.6	271	16 Q98BX7	Q98BX7 rhizobium 1
25	31	75.6	271	16 Q92NE0	Q92NE0 rhizobium m
26	31	75.6	302	5 Q9W335	Q9W335 drosophila
27	31	75.6	343	2 Q53025	Q53025 rhodococcus
28	31	75.6	347	16 Q9KGN0	Q9KGN0 bacillus ha
29	31	75.6	386	12 Q9DXY2	Q9DXY2 plutella xy
30	30	73.2	52	9 Q9B036	Q9B036 mycobacteri
31	30	73.2	99	11 Q9ERU1	Q9ERU1 rattus norv
32	30	73.2	104	4 Q9NRF2	Q9NRF2 homo sapien
33	30	73.2	108	17 Q9FYF0	Q9FYF0 aeropyrum p
34	30	73.2	146	13 Q9DF33	Q9DF33 ophiophagus
35	30	73.2	222	5 Q9V8H1	Q9V8H1 drosophila
36	30	73.2	222	5 Q9V8H0	Q9V8H0 drosophila
37	30	73.2	238	5 Q9Y098	Q9Y098 periplaneta
38	30	73.2	240	5 Q9V8G5	Q9V8G5 drosophila
39	30	73.2	283	16 Q9AAJ7	Q9AAJ7 caulobacter
40	30	73.2	331	16 Q9HYE6	Q9HYE6 pseudomonas
41	30	73.2	388	16 Q9RSN6	Q9RSN6 deinococcus
42	30	73.2	410	2 Q9XEV7	Q9XEV7 streptomyce
43	30	73.2	425	17 Q59010	Q59010 pyrococcus
44	30	73.2	452	10 Q9ZQ00	Q9ZQ00 arabidopsis
45	30	73.2	466	6 Q95KH6	Q95KH6 macaca fasc

ALIGNMENTS

RESULT 1
Q96TH4 ID Q96TH4 PRELIMINARY: PRT: 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TARA-AMYLASE A (EC 3.2.1.11) (ALPHA-AMYLASE).
CN AMYA OR AMY1
OS Aspergillus oryzae, and
OS Aspergillus flauus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062, 5059;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A.oryzae; STRAIN=R1840;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
RA Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator
gene, amyR, involved in the amylolytic gene expression in Aspergillus
oryzae.";
RT Biosci. Biotechnol. Biochem. 64:816-827(2000).
RL [2]
RN SEQUENCE FROM N.A.
RP SPECIES=A.flavus; STRAIN=86-10D;
RC Fakhoury A.M., Moloshuk C.P.;
RA "Amyl, the alpha-amylase gene of Aspergillus flavus: Involvement in
afatoxin biosynthesis in maize kernels.";
RL Phytopathology 89:908-914(1999).
DR EMBL; AB021876; BAA9703.1; -;
DR EMBL; AF139925; BAA9703.1; -;
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match 100.0%; Score 41; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
 |||||
 Db 449 YTAGOOLT 456

RESULT 2

ID 09KND8 PRELIMINARY; PRT; 849 AA.
 AC 09KND8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CHITINASE.
 GN VCA0027.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 NCBI_TaxID=66;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., McKalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae.";
 RT Nature 406:477-483(2000).
 CC -1. SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 HYDROLASES).
 DR EMBL: AE004346; AAF95941.1; -.
 DR HSSP: P07254; ICTN.
 DR TIGR: VCA0027; -.
 DR InterPro: IPR001579; Chitinase_2.
 DR InterPro: IPR003610; Chitin_bind3.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR InterPro: IPR000601; PKD domain.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR Pfam: PF00801; PKD; 2.
 DR SMART; SM00495; ChcbD3; 1.
 DR SMART; SM00089; PKD; 2.
 DR PROSITE; PS01095; CHITINASE_18; 1.
 DR Complete proteome: Glycosidase; Hydrolase.
 KW Complete proteome: Glycosidase; Hydrolase.
 SQ SEQUENCE 849 AA; 90314 MW; DB164D55673F6D34 CRC64;

Query Match 82.9%; Score 34; DB 16; Length 849;
 Best Local Similarity 75.0%; Pred. No. 57;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
 |||||
 Db 805 YTAGOOLT 812

RESULT 3

ID 0982J7 PRELIMINARY; PRT; 420 AA.
 AC 0982J7;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE ABC TRANSPORTER, PERIPLASMIC PROTEIN.
 GN MUR9034.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Plasmid pMA.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 NCBI_TaxID=381;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFE303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RT DNA Res. 7:331-338(2000).
 DR EMBL: AP003015; BAB54459.1; -.
 KW Plasmid; Complete proteome.
 SQ SEQUENCE 420 AA; 45026 MW; 95E3EF93E2CB093 CRC64;

Query Match 80.5%; Score 33; DB 16; Length 420;
 Best Local Similarity 75.0%; Pred. No. 45;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
 |||||
 Db 83 YTAGOOLT 90

RESULT 4

ID 09YWK5 PRELIMINARY; PRT; 192 AA.
 AC 09YWK5;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE HYPOTHETICAL 22.7 KDA PROTEIN (FRAGMENT).
 OS Buzura suppressaria nuclear polyhedrosis virus (BsnPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 NCBI_TaxID=74320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99036048; PubMed=9820162;
 RA Hu Z.H., Arif B.M., Jin F., Martens J.W.M., Chen X.W., Sun J.S.,
 RA Zuidema D., Goldbach R.W., Vlak J.M.;
 RT "Distinct gene arrangement in the Buzura suppressaria single-
 RT nucleocapsid nucleopolyhedrovirus genome.";
 RT J. Gen. Virol. 79:2841-2851(1998).
 DR EMBL: AF058929; AAC77811.1; -.
 KW Hypothetical protein.
 FT NON_TER 192
 SQ SEQUENCE 192 AA; 22713 MW; D974F7B02895DC8E CRC64;

Query Match 78.0%; Score 32; DB 12; Length 192;
 Best Local Similarity 75.0%; Pred. No. 32;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
 |||||
 Db 141 YTAGOOLT 148

RESULT 5

ID 09GPL6 PRELIMINARY; PRT; 211 AA.
 AC 09GPL6;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE 1-10.
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 OC Anophelinae.

OX NCBI_TaxID=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ZAN/7.
 RA Orrelli F., Ranson H., Jensen B., Rossiter L., Wang X., Roth C.,
 RA Collins F.H., Hemingway J.;
 RT "Identification of a novel class of insect glutathione S-transferases
 RT associated with DDT resistance in the malaria vector *Anopheles*
 RT *gambiae*.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF316638; AAC5166.1; -
 DR InterPro: IPR004046; GST_C.
 DR InterPro: IPR004045; GST_N.
 KW Transferase.
 SQ SEQUENCE 211 AA; 24439 MW; F37363CE18ADE71 CRC64;

Query Match 78.0%; Score 32; DB 5; Length 211;
 Best Local Similarity 75.0%; Pred. No. 36;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
 |||||
 DB 144 YVAGQKLT 151

RESULT 6
 ID 099H12 PRELIMINARY; PRT; 333 AA.
 AC 099H12;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE P47.
 OS *Helicoverpa armigera* nucleopolyhedrovirus G4.
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 OX NCBI_TaxID=148363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Deng F., Chen X., Vlak J.M., Arif B.M., Hu Z.;
 RT "Sequence analysis of the gp17 gene of *Heliothis armigera* single-
 RT nucleocapsid nucleopolyhedrovirus.";
 RL Zhongguo Bingdaxue 15:35-42(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wang H., Hu Z., Sun X., Vlak J.M., Chen X.;
 RT "Sequence analysis of the iap3 gene of *Heliothis armigera* single-
 RT nucleocapsid nucleopolyhedrovirus.";
 RL Zhongguo Bingdaxue 15:43-49(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21078302; PubMed=11210934;
 RA Wang H., Chen X., Wang H., Arif B.M., Vlak J.M., Hu Z.;
 RT "Nucleotide sequence and transcriptional analysis of a putative basic
 RT DNA-binding protein of *Helicoverpa armigera* polyhedrovirus.";
 RL Virus Genes 22:113-120(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21064569; PubMed=11125177;
 RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
 RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;
 RT "The sequence of the *Helicoverpa armigera* single-nucleocapsid
 RT nucleopolyhedrovirus genome.";
 RL J. Gen. Virol. 82:241-257(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
 RA Peters S., Zuidema D., Lankhorst R.K., Vlak J.M., Hu Z.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF211059; AAC53778.1; -
 SQ SEQUENCE 333 AA; 38963 MW; B8EBDAF391A40285 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 333;
 Best Local Similarity 75.0%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
 |||||
 DB 161 YTLGQQLS 168

RESULT 7
 ID 09YMS6 PRELIMINARY; PRT; 390 AA.
 AC 09YMS6;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE LD-P47.
 OS *Lymantria dispar* multicapsid nuclear polyhedrosis virus (LdMNPV).
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10449;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99124785; PubMed=9887315;
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J.M., Rohmann G.F.;
 RT "Sequence and analysis of the genome of a baculovirus pathogenic for
 RT *Lymantria dispar*.";
 RL Virology 253:17-34(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kuzio J.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF081810; AAC70233.1; -
 SQ SEQUENCE 390 AA; 45535 MW; 1A982632B037C552 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 390;
 Best Local Similarity 75.0%; Pred. No. 69;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
 |||||
 DB 140 YTLGQQLS 147

RESULT 8
 ID 09PYW5 PRELIMINARY; PRT; 394 AA.
 AC 09PYW5;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ORF78.
 OS *Xestia c-nigrum* granulosis virus (XngV) (*Xestia c-nigrum*
 OS *granulovirus*).
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
 OX NCBI_TaxID=51677;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99434230; PubMed=10502508;
 RA Hayakawa T., Ko R., Okano K., Seong S.I., Goto C., Maeda S.;
 RT "Sequence analysis of the *Xestia c-nigrum* granulovirus genome.";
 RL Virology 262:277-297(1999).
 DR EMBL: AF162221; AAF05192.1; -
 SQ SEQUENCE 394 AA; 46157 MW; 4C598BFCB5235DBF CRC64;

Query Match 78.0%; Score 32; DB 12; Length 394;
 Best Local Similarity 75.0%; Pred. No. 70;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 || |||||
 Db 141 YTLGOQLS 148

RESULT 9
 Q9IML8 PRELIMINARY; PRT; 399 AA.

AC Q9IML8; MEDLINE=97329351; PubMed=9185864;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE P47.
 GN P47.
 OS Choristoneura fumiferana nuclear polyhedrosis virus (cfMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10448;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EC-1;
 RX MEDLINE=20276145; PubMed=10814576;
 RA Lapointe R., Back D.W., Ding Q., Carstens E.B.;
 RT "Identification and molecular characterization of the choristoneura
 fumiferana multicausid nucleopolyhedrovirus genomic region encoding
 the regulatory genes pkip, p47, lef-12, and gta.";
 RT Virology 271:109-121(2000).
 RL EMBL: AF177329; AAF64092.1; -
 DR EMBL: AF177329; AAF64092.1; -
 SQ SEQUENCE 399 AA; 46323 MW; 744600382F7DC942 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 399;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 || |||||
 Db 148 YTLGOQLS 155

RESULT 10
 Q92407 PRELIMINARY; PRT; 399 AA.

AC Q92407; MEDLINE=97329351; PubMed=9185864;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE P47-ACMNPV ORF40.
 GN P47.
 OS Bombyx mori nuclear polyhedrosis virus (BmNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10458;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;
 RX MEDLINE=97329351; PubMed=9185864;
 RA Kamita S.G., Maeda S.;
 RT "Sequencing of the putative DNA helicase-encoding gene of the Bombyx
 mori nuclear polyhedrosis virus and fine-mapping of a region involved
 in host range expansion.";
 RT Gene 190:173-179(1997).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=T3;
 RC STRAIN=T3;
 RX MEDLINE=99281911; PubMed=10355780;
 RA Gomi S., Majima K., Maeda S.;
 RT "Sequence analysis of the genome of Bombyx mori
 nucleopolyhedrovirus.";
 RT J. Gen. Virol. 80:1323-1337(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;

RA Maeda S.;
 RL Submitted (May-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;
 RA Gomi S.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U3180; AAC63715.1; -
 SQ SEQUENCE 399 AA; 47260 MW; 69114411B659C63 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 399;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 || |||||
 Db 144 YTLGOQLS 151

RESULT 11
 Q9IGT4 PRELIMINARY; PRT; 399 AA.

AC Q9IGT4; MEDLINE=20036646; PubMed=10567663;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE P47.
 GN P47.
 OS Epiphyas postvittana nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=70600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hyink O., Dellow R.A., Olsen M., Caradoc-Davies K.M.B., Drake K.,
 RA Ward V.K.;
 RT "The complete sequence of the Epiphyas postvittana
 nucleopolyhedrovirus genome.";
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY043265; AAK85601.1; -
 SQ SEQUENCE 399 AA; 46410 MW; 6495E5250ABFAF11 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 399;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 || |||||
 Db 148 YTLGOQLS 155

RESULT 12
 Q9J822 PRELIMINARY; PRT; 400 AA.

AC Q9J822; MEDLINE=20036646; PubMed=10567663;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ORF115 P47.
 GN Spodoptera exigua nucleopolyhedrovirus.
 OS Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;
 RX MEDLINE=20036646; PubMed=10567663;
 RA IJbel W.F., van Strien E.A., Heidens J.G., Broer R., Zuidema D.,
 RA Goldbach R.W., Vlak J.M.;
 RT "Sequence and organization of the spodoptera exigua multicausid
 nucleopolyhedrovirus genome.";
 RT J. Gen. Virol. 80:3289-3304(1999).
 RL

RN [2]
 RN SEQUENCE FROM N.A.
 RA Jukei W.F.J., van Strien E.A., Helkens J.G.M., Broer R., Zuidema D.,
 RA Goldbach R.W., Vlak J.M.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF169823; AAF33644.1; -
 SO SEQUENCE 400 AA; 46896 MW; 0373F30953553F5B CRC64;

Query Match 78.0%; Score 32; DB 12; Length 400;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YTAGOOLT 8
 Db 144 YTLGQOLS 151

RESULT 13
 Q91BX7 PRELIMINARY; PRT; 412 AA.
 AC Q91BX7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE P47.
 OS Helicoverpa armigera nuclear polyhedrosis virus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_Taxid=51313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C1;
 RA Zhang C.X., Wu J.C.;
 RT "Genome structure and the p10 gene of the Helicoverpa armigera
 RT nucleopolyhedrovirus.";
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 33:179-184(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C1;
 RA Zhang C.X., Jin W.R.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF303045; AAK96288.1; -
 SO SEQUENCE 412 AA; 48091 MW; DBA6452E9159AA85 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 412;
 Best Local Similarity 75.0%; Pred. No. 74;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
 Db 161 YTLGQOLS 168

RESULT 14
 Q91B15 PRELIMINARY; PRT; 422 AA.
 AC Q91B15;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE TRANSCRIPTION REGULATOR P47.
 OS Spodoptera litura nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_Taxid=46242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=G2;
 RX MDLLine=21425398; PubMed=11531416;
 RA Pang Y., Yu J., Wang L., Hu X., Bao W., Li G., Chen C., Han H., Hu S.,
 RA Yang H.;

RT *Sequence Analysis of the Spodoptera litura Multicapsid
 RT Nucleopolyhedrovirus Genome.";
 RL Virology 287:391-404(2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=G2;
 RA Yu J., Wang L., Hu X., Pang Y.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF325155; AAL01722.1; -
 SO SEQUENCE 422 AA; 49358 MW; 64D30486B1522621 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 422;
 Best Local Similarity 75.0%; Pred. No. 75;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YTAGOOLT 8
 Db 152 YTLGQOLS 159

RESULT 15
 Q91EY7 PRELIMINARY; PRT; 460 AA.
 AC Q91EY7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ORF68 P47.
 GN ORF68 P47.
 OS Cydia pomonella granulosis virus (CPGV) (Cydia pomonella
 OS granulovirus).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
 OX NCBI_Taxid=28289;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=9318168; PubMed=8445726;
 RA Crook N.E., Clem R.J., Miller L.K.;
 RT "An apoptosis-inhibiting baculovirus gene with a zinc finger-like
 RT motif.";
 RL J. Virol. 67:2168-2174(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=96207404; PubMed=8615018;
 RA Theilmann D.A., Chantler J.K., Stewart S., Flipsen H.T., Vlak J.M.,
 RA Crook N.E.;
 RT "Characterization of a highly conserved baculovirus structural protein
 RT that is specific for occlusion-derived virions.";
 RL Virology 218:148-158(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=97380577; PubMed=9237352;
 RA Kang W., Crook N.E., Winstanley D., O'Reilly D.R.;
 RT "Complete sequence and transposon mutagenesis of the BamHI J fragment
 RT of Cydia pomonella granulosis virus.";
 RL Virus Genes 14:131-136(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=98418511; PubMed=9747739;
 RA Kang W., Tristem M., Maeda S., Crook N.E., O'Reilly D.R.;
 RT "Identification and characterization of the Cydia pomonella
 RT granulovirus cathepsin and chitinase genes.";
 RL J. Gen. Virol. 79:2283-2292(1998).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RA Luque T., Finch R., Crook N., O'Reilly D.R., Winstanley D.;
 RT "The complete sequence of the Cydia pomonella granulovirus genome.";
 RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: U53466; AAK70728.1; -
 SO SEQUENCE 460 AA; 54126 MW; 8219FA11B7DD1A6B CRC64;

Query Match 78.0%; Score 32; DB 12; Length 460;
 Best Local Similarity 75.0%; Pred. No. 83;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YTAGGQOLT 8
 || |||||
 Db 204 YTLGQOLS 211

Search completed: June 13, 2002, 09:11:17
 Job time: 1062 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:54 ; Search time 38.1 Seconds

(without alignments)
8.130 Million cell updates/sec

Title: US-09-710-339-2_COPY_448_455

Perfect score: 41

Sequence: 1 YTAGQQLT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	498	1 AMYA_ASPAW	Q02905 aspergillus
2	41	100.0	499	1 AMYA_ASPOR	P10529 aspergillus
3	41	100.0	499	1 AMYA_ASPAW	Q02906 aspergillus
4	41	100.0	499	1 AMY_ASPSH	P30292 aspergillus
5	33	80.5	370	1 NREF_HAELN	P44942 haemophilus
6	32	78.0	399	1 VP47_NPYOP	O10300 orgyia pseu
7	32	78.0	401	1 VP47_NPVAC	P34051 autographa
8	31	75.6	215	1 GT23_DROME	Q9vg98 drosophila
9	31	75.6	215	1 GT23_DROME	Q9vg96 drosophila
10	31	75.6	226	1 YUG9_VITCO	Q9kkk6 streptomyce
11	31	75.6	282	1 YUG9_VITCO	Q9kkyl vibrio chol
12	31	75.6	442	1 NX1B_HUMAN	P58400 homo sapien
13	31	75.6	554	1 CX1A_PARDE	P08305 paracoccus
14	31	75.6	558	1 CX1B_PARDE	P98002 paracoccus
15	31	75.6	699	1 CH11_BACCI	P20533 bacillus ci
16	31	75.6	1477	1 NX1A_HUMAN	Q9ulb1 homo sapien
17	31	75.6	5179	1 MOC2_HUMAN	Q02817 homo sapien
18	30	73.2	390	1 DACD_SALTY	P37604 salmonella
19	30	73.2	449	1 GUNA_CELFI	P07984 cellulomona
20	30	73.2	473	1 NORB_PSEST	P98008 pseudomonas
21	29	70.7	105	1 CUI2_HYACE	P45588 hyalophora
22	29	70.7	270	1 CB12_LYCES	P10708 lycopersico
23	29	70.7	270	1 CB12_LYCES	P10708 lycopersico
24	29	70.7	366	1 TRPD_MYCLE	P13869 petunia hyb
25	29	70.7	388	1 DACD_ECOLI	P09581 mycobacteri
26	29	70.7	466	1 NORB_PSEAE	P33013 escherichia
27	29	70.7	512	1 YW09_MYCMU	O59647 pseudomonas
28	29	70.7	551	1 LIND_ECOLI	O10398 mycobacteri
29	29	70.7	565	1 COX1_RHOSH	P33231 escherichia
30	29	70.7	728	1 EF2_HALHA	P14823 rhodobacter
31	29	70.7	735	1 ABC8_HUMAN	Q9nut2 homo sapien
32	29	70.7	922	1 CLPA_PEA	P5100 plism sativ
33	29	70.7	923	1 CLAB_LYCES	P15452 lycopersico

34	29	70.7	926	1 CLAB_LYCES	P15451 lycopersico
35	29	70.7	1024	1 CARG_HUMAN	Q9npp4 homo sapien
36	29	70.7	1231	1 CFAB_HUMAN	P08603 homo sapien
37	29	70.7	1328	1 YMD9_YEAST	P03434 saccharomyc
38	29	70.7	1328	1 YMD9_YEAST	O04214 saccharomyc
39	29	70.7	1375	1 NID2_HUMAN	O14112 homo sapien
40	29	70.7	1463	1 PA2R_BOVIN	P49259 bos taurus
41	28	68.3	87	1 V476_MYCLE	Q9c843 mycobacteri
42	28	68.3	129	1 R13_TERPS	P11252 yersinia ps
43	28	68.3	133	1 REV_CAEYC	P33460 caprine art
44	28	68.3	210	1 GRT3_MUSDO	P46432 musca domes
45	28	68.3	217	1 RL3_MYCBO	O06044 mycobacteri

ALIGNMENTS

RESULT	ID	AMYA_ASPAW	STANDARD	PRT	498 AA.
AC	Q02905	AMYA_ASPAW	STANDARD	PRT	498 AA.
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-OCT-1996	(Rel. 34, Last annotation update)			
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan				
DE	glucanohydrolase A).				
OS	AMYA.				
CC	Aspergillus awamori.				
CC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
CC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_Taxid=105351;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=UK143F;				
RX	MEDLINE=90254827; Pubmed=2340591;				
RA	Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,				
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,				
RA	Berka R.M.;				
RT	"Cloning, characterization, and expression of two alpha-amylase genes				
RT	from Aspergillus niger var. awamori.";				
RL	Curr. Genet. 17:203-212(1990).				
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic				
CC	linkages in oligosaccharides and polysaccharides.				
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO				
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-slb.ch/announce/				
CC	or send an email to license@isb-slb.ch).				
DR	EMBL; X52755; CAA36966.1; -				
DR	HSSP; P10529; TTA.				
DR	InterPro: IPR000461; Alpha-amylase.				
DR	Pfam; PF00128; alpha-amylase; 1.				
KW	Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;				
KW	Glycoprotein; Signal; Multigene family.				
FT	SIGNAL	1	21	POTENTIAL.	
FT	CHAIN	22	498	ALPHA-AMYLASE A.	
FT	ACT_SITE	227	227	BY SIMILARITY.	
FT	ACT_SITE	231	231	BY SIMILARITY.	
FT	ACT_SITE	251	251	BY SIMILARITY.	
FT	ACT_SITE	251	251	BY SIMILARITY.	
FT	DISULFID	318	318	BY SIMILARITY.	
FT	DISULFID	51	59	BY SIMILARITY.	
FT	DISULFID	171	185	BY SIMILARITY.	
FT	DISULFID	261	304	BY SIMILARITY.	
FT	DISULFID	461	496	BY SIMILARITY.	
FT	CARBOHYD	218	218	N-LINKED (GLCNAC...) (POTENTIAL).	
SQ	SEQUENCE	498 AA;	54880 MW;	7658511BC01A8A01 CRC64;	

Query Match 100.0%; Score 41; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
 |||||
 Db 449 YTAGOOLT 456

RESULT 2

AMTA_ASPOR STANDARD; PRT; 499 AA.
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 DE (1,4-alpha-D-glucan glucanohydrolase).
 GN AMT1 AND AMT2 AND AMT3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 63303;
 RX MEDLINE=89237897; Pubmed=2785629;
 RA Wiersel S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT intron-exon organization.";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89378767; Pubmed=2789162;
 RA Genes M.J., Dove M.J., Seligy V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns.";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae.";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90128276; Pubmed=2612911;
 RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
 RA Ueda S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes.";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A.";
 RL Proc. Jpn. Acad. B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RX MEDLINE=74001521; Pubmed=4733850;
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RT A with trypsin and chymotrypsin.";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RT Proc. Jpn. Acad. B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; Pubmed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution.";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84212370; Pubmed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A.";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; Pubmed=9283074;
 RA Brozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution.";
 RL Biochemistry 36:10837-10845(1997).
 CC -I CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I COPFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -I SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04548; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: JT0466; JT0466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR GlycosultedB: P10529; -;
 DR InterPro: IPR000461; Alpha.amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3d-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT CONFLICT 93 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTID-CAR.000125.
 O -> R (IN AMY3).
 F -> L (IN AMY3).
 T* -> DC (IN REF. 5).
 O -> T (IN REF. 5).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 G -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).

FT CONFLICT 370 370 L -> A (IN REF. 4).
 FT CONFLICT 406 409 WPLY -> PYI (IN REF. 5).
 FT CONFLICT 448 448 G -> S (IN REF. 5).
 FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).
 FT HELIX 24 27
 FT STRAND 28 29
 FT STRAND 32 35
 FT HELIX 37 40
 FT STRAND 42 42
 FT STRAND 43 44
 FT HELIX 53 55
 FT STRAND 61 61
 FT HELIX 63 67
 FT TURN 68 69
 FT HELIX 70 74
 FT TURN 75 77
 FT STRAND 80 83
 FT STRAND 87 89
 FT STRAND 94 94
 FT TURN 95 96
 FT STRAND 97 97
 FT TURN 100 101
 FT STRAND 105 111
 FT TURN 113 115
 FT TURN 118 119
 FT HELIX 120 130
 FT TURN 131 132
 FT STRAND 134 139
 FT STRAND 143 143
 FT STRAND 146 146
 FT HELIX 150 152
 FT TURN 155 156
 FT STRAND 159 159
 FT HELIX 164 166
 FT STRAND 167 167
 FT STRAND 172 172
 FT HELIX 179 184
 FT STRAND 186 188
 FT STRAND 193 194
 FT STRAND 196 197
 FT TURN 199 200
 FT HELIX 202 219
 FT TURN 220 220
 FT STRAND 223 226
 FT TURN 227 228
 FT HELIX 229 231
 FT TURN 234 235
 FT HELIX 236 244
 FT STRAND 247 250
 FT HELIX 257 265
 FT TURN 266 266
 FT STRAND 269 271
 FT HELIX 273 283
 FT TURN 286 287
 FT HELIX 307 309
 FT STRAND 311 312
 FT TURN 317 318
 FT HELIX 322 325
 FT STRAND 341 341
 FT STRAND 345 349
 FT TURN 350 351
 FT HELIX 352 354
 FT TURN 355 355
 FT TURN 361 364
 FT HELIX 368 371
 FT TURN 372 372

Query Match 100.0%; Score 41; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8

Db 449 YTAGOQLT 456
 RESULT 3
 ID AMYB_ASPAW STANDARD; PRT; 499 AA.
 AC 002906;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Alpha-amyase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 glucanohydrolase B).
 GN AMYB.
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=105351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OVK13F;
 RX MEDLINE=90254827; PubMed=2340591;
 RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
 RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
 RA Berka R.M.;
 RT "Cloning, characterization, and expression of two alpha-amyase genes
 from Aspergillus niger var. awamori.";
 RL Curr. Genet. 17:203-212(1990).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 linkages in oligosaccharides and polysaccharides.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 KNOWN AS THE ALPHA-AMYLASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X52756; CAA36967.1; -
 DR HSSP; P10529; YTA.
 DR InterPro; IPR000461; Alpha-amyase.
 DR Pfam; PF00128; alpha-amyase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 SO SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;

Query Match 100.0%; Score 41; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8

Db 449 YTAGOQLT 456
 RESULT 4
 AMY_ASPSW

```

ID  AMY_ASPIH  STANDARD;  PRT;  499 AA.
AC  P30292;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY.
OS  Aspergillus shirousami.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC  Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX  NCBI_TaxID=5070;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92323146; PubMed=1368777;
RX  Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT  "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
RT  expression in Saccharomyces cerevisiae."
RL  Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC  -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D10461; BAA01255.1; -
DR  PIR: J50663; J50663.
DR  HSP: P10529; 7TAA.
DR  InterPro: IPR000461; Alpha_Amylase.
DR  Pfam: PF00128; alpha-amylase; 1.
KW  Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KW  Glycoprotein; Signal.
FT  SIGNAL 1 21
FT  CHAIN 22 499 ALPHA-AMYLASE.
FT  ACT_SITE 227 227
FT  ACT_SITE 231 231
FT  ACT_SITE 251 251 BY SIMILARITY.
FT  ACT_SITE 318 318 BY SIMILARITY.
FT  ACT_SITE 51 59 BY SIMILARITY.
FT  DISULFID 171 185 BY SIMILARITY.
FT  DISULFID 261 304 BY SIMILARITY.
FT  DISULFID 461 496 BY SIMILARITY.
FT  CARBOHYD 218 218 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ  SEQUENCE 499 AA; 54852 MW; 1FB7AE50DA01C03F CRC64;

Query Match 100.0%; Score 41; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
DB 449 YTAGOOLT 456

```

```

OC  Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC  Haemophilus.
OX  NCBI_TaxID=727;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=RD / KW20 / ATCC 51907;
RX  MEDLINE=95350630; PubMed=7542800;
RA  Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA  Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA  McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA  Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA  Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA  Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA  Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
RA  Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA  Venter J.C.;
RT  Whole-genome random sequencing and assembly of Haemophilus
RT  influenzae Rd.
RL  Science 269:496-512(1995).
CC  -1- FUNCTION: MAY BE REQUIRED FOR THE BIOGENESIS OF C-TYPE
CC  CYTOCHROMES. POSSIBLE SUBUNIT OF A HEME LYASE (BY SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC  -1- SIMILARITY: BELONGS TO THE CCMH/CYCL/CCL2/NRFF FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: U32775; AAC22592.1; ALT_INIT.
DR  TIGR: H10934; -
DR  InterPro: IPR001440; TPR.
DR  Pfam: PF00515; TPR; 1.
KW  Cytochrome c-type biogenesis; Periplasmic; Heme; Signal;
KW  Complete proteome.
FT  SIGNAL 1 21 POTENTIAL.
FT  CHAIN 22 370 CYTOCHROME C-TYPE BIOGENESIS PROTEIN
FT  BINDING 47 47 NRFF.
FT  BINDING 50 50 HEME (COVALENT) (POTENTIAL).
FT  BINDING 50 50 HEME (COVALENT) (POTENTIAL).
SQ  SEQUENCE 370 AA; 42161 MW; 4023B680D7EA9480 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 370;
Best Local Similarity 75.0%; Pred. No. 6.6;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
DB 287 YTAGOOLT 294

```

```

RESULT 5
NRFF_HAEIN  STANDARD;  PRT;  370 AA.
AC  P44942;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Cytochrome c-type biogenesis protein nrff precursor.
GN  NRFF OR H10934.
OS  Haemophilus influenzae.

```

```

RESULT 6
VP47_NPYOP  STANDARD;  PRT;  399 AA.
ID  VP47_NPYOP  STANDARD;  PRT;  399 AA.
AC  O10300;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  Viral transcription regulator P47.
GN  P47.
OS  Orgyia pseudotsugata multicausid polyhedrosis virus (OPMNFV).
OC  Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC  Nucleopolyhedrovirus.
OX  NCBI_TaxID=164623;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=97271300; PubMed=9126251;
RX  Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RX  Rohrmann G.F.;

```

```

RT      "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RT      polyhedrosis virus genome."
RL      Virology 229:381-399(1997).
CC      -1- FUNCTION: IT IS INVOLVED IN REGULATING VIRAL TRANSCRIPTION AT
CC      LATE TIMES POSTINFECTION (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: NUCLEUS OF INFECTED CELLS (BY SIMILARITY).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U75930; AAC59044.1; -
KW      Transcription regulation; Nuclear protein.
SQ      SEQUENCE 399 AA; 46357 MW; A5476EDCF83225C0 CRC64;

Query Match      78.08; Score 32; DB 1; Length 399;
Best Local Similarity 75.08; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy      1 YTAGOQLT 8
        |||||
Db      148 YTIQGOLS 155

RESULT 7
VP47_NPVAC STANDARD; PRT; 401 AA.
AC      P34051;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      01-NOV-1995 (Rel. 32, Last annotation update)
DE      Viral transcription regulator p47.
GN      p47.
OS      Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC      Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC      Nucleopolydovirus.
OX      NCBI_TaxID=46015;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=HR3;
RX      MEDLINE=93233214; PubMed=8474157;
RA      Carstens E.B., Lu A.L., Chan H.L.B.;
RT      "Sequence, transcriptional mapping, and overexpression of p47, a
RT      baculovirus gene regulating late gene expression.",
RL      J. Virol. 67:2513-2520(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C6;
RX      MEDLINE=94303173; PubMed=8030224;
RA      Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT      "The complete DNA sequence of Autographa californica nuclear
RT      polyhedrosis virus."
RL      Virology 202:586-605(1994).
CC      -1- FUNCTION: IT IS INVOLVED IN REGULATING VIRAL TRANSCRIPTION AT
CC      LATE TIMES POSTINFECTION.
CC      -1- SUBCELLULAR LOCATION: NUCLEUS OF INFECTED CELLS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; L07878; AAA16859.1; -
DR      EMBL; L22858; AAA66670.1; -
DR      PIR; B45684; B45684.

```

```

KW      Transcription regulation; Nuclear protein.
FT      VARIANT 131 131 V -> M (TEMPERATURE SENSITIVE MUTANT).
SQ      SEQUENCE 401 AA; 47530 MW; 6E1818027042DC3A CRC64;

Query Match      78.08; Score 32; DB 1; Length 401;
Best Local Similarity 75.08; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy      1 YTAGOQLT 8
        |||||
Db      144 YTIQGOLS 151

RESULT 8
GT21_DROME STANDARD; PRT; 215 AA.
ID      GT21_DROME
AC      Q9VG98; Q9T9X1;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Glutathione S-transferase D21 (EC 2.5.1.18) (DMGST21).
GN      GSTD21 OR CG4181.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC      Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93252851; PubMed=7683659;
RA      Young Y.-P.S., Heieh T.-S., Tu C.-P.D.;
RT      "The glutathione S-transferase D genes, A divergently organized,
RT      intronless gene family in drosophila melanogaster.",
RL      J. Biol. Chem. 268:9737-9746(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BERKELEY;
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman U.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA      Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballow R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Bortova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA      Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA      Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA      Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA      Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA      Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kelthum K.A.,
RA      Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA      Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA      Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T.,
RA      Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA      Swidski R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA      Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA      Ye J., Yeh K.-F., Zaveri J.S., Zhang W., Zhang G., Zhao Q., Zheng L.,

```

```

RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster.";
RL  Science 287:2185-2195(2000)
CC  -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF
CC  EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES (BY
CC  SIMILARITY).
CC  -1- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC  -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. THETA FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: M97702; -; NOT_ANNOTATED_CDS.
DR  EMBL: AE003695; AAF54787.1; -
DR  FlyBase: FBgn0010038; Gstd21.
DR  InterPro: IPR004046; GST_C.
DR  InterPro: IPR004045; GST_N.
DR  Pfam: PF00043; GST_C; 1.
DR  Pfam: PF02798; GST_N; 1.
DR  Transfaser: Multigene family.
KM  "The genome sequence of Drosophila melanogaster."
SQ  SEQUENCE 215 AA; 24535 MW; 221A961BBF0E5B5 CRC64;

Query Match          75.6%; Score 31; DB 1; Length 215;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1 YTAGOOLT 8
    1 11111
Db  146 YVAGDOLT 153

RESULT  9
G723_DROME STANDARD; PRT; 215 AA.
ID  G723_DROME
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Glutathione S-transferase D23 (EC 2.5.1.18) (DMGST23).
GN  GSTD23 OR CG11512.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxId=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  MEDLINE-93252851; PubMed-7683659;
RA  Young Y.-P.S., Hsieh T.-S., Yu C.-P.D.;
RT  "The glutathione S-transferase D genes. A divergently organized,
RT  intronless gene family in Drosophila melanogaster."
RL  J. Biol. Chem. 268:9737-9746(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  STRAIN-BERKELEY;
RX  MEDLINE-20196006; PubMed-10731132.
RA  Adams M.D., Celinker S.E., Holt R.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA  Brandon R.C., Rogers Y.-H.C., Blazer R.C., Champe M., Pfeiffer B.D.,
RA  Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA  Adair J.F., Abdayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA  Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

```

```

RA  Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA  Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Chert J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays R.D., Dew I., Dietz S.M.,
RA  Doonan K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA  Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA  Jaisli M., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Lako P., Lei Y., Levitsky A.A., Li U., Li Z., Liang Y., Lin X.,
RA  Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA  Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA  Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA  Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA  Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster."
RL  Science 287:2185-2195(2000)
CC  -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF
CC  EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES (BY
CC  SIMILARITY).
CC  -1- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC  -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. THETA FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: M97702; -; NOT_ANNOTATED_CDS.
DR  EMBL: AE003695; AAF54789.1; -
DR  FlyBase: FBgn0010040; Gstd23.
DR  InterPro: IPR004046; GST_C.
DR  InterPro: IPR004045; GST_N.
DR  Pfam: PF00043; GST_C; 1.
DR  Pfam: PF02798; GST_N; 1.
DR  Transfaser: Multigene family.
KM  "The genome sequence of Drosophila melanogaster."
SQ  SEQUENCE 215 AA; 24736 MW; 9556E0A1A54BBE1F CRC64;

Query Match          75.6%; Score 31; DB 1; Length 215;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1 YTAGOOLT 8
    1 11111
Db  146 YVAGDOLT 153

RESULT  10
P1RO_STRCO STANDARD; PRT; 226 AA.
ID  P1RO_STRCO
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Phosphoribosyltransferase I (EC 6.3.5.3) (PGAM

```



```
DE synthase 1).
GN PURO OR SCD25.14.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D., Thomson N.R., Parkhill J., Barrell B.G.,
RA Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + 5'-phosphoribosylformylglycinamide + L-
CC glutamine + H(2)O = ADP + phosphate + 5'-
CC phosphoribosylformylglycinamide + L-glutamate.
CC -1- PATHWAY: DE NOVO PURINE BIOSYNTHESIS; FOURTH STEP.
CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, PURO AND PURL.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity)
CC -1- SIMILARITY: BELONGS TO TYPE-1 GLUTAMINE AMIDOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AL118514; CAB56358.1; -
DR InterPro: IPR000991; GATase.1.
DR PROSITE: PS00442; GATASE_TYPE_I; 1.
FM Purine biosynthesis; Ligase; Glutamine amidotransferase.
KT ACT_SITE 87 87 GATASE (By similarity).
SQ SEQUENCE 226 AA; 24528 MW; 4744E8706CD39D23 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 226;
Best Local Similarity 71.4%; Pred. NO. 11;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOL 7
Db 131 YTAGOEI 137
|||||:
131 YTAGOEI 137

RESULT 11
ID Y1G9_VIBCH STANDARD: PRT; 282 AA.
AC Q9KKY1.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein VCA0969.
GN VCA0969.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision: Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Radoqi I., Sellers P.,
RA McDonald L., Uitterlind T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- SIMILARITY: BELONGS TO THE PIRIN FAMILY.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AE004423; AAF96865.1; -
DR TIGR, VCA0969; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 282 AA; 31324 MW; 5010C3933BE13197 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 282;
Best Local Similarity 85.7%; Pred. NO. 14;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YAGOOIT 8
Db 226 YAGOOIT 232
|||||:
226 YAGOOIT 232

RESULT 12
ID NX1B_HUMAN STANDARD: PRT; 442 AA.
AC P58400.
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neurexin 1-beta precursor (Neurexin I-beta).
GN NRXN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Yohikawacatarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99118874; PubMed=9921901;
RA Kleiderlein J.U., Nilsson P.E., Jessee J., Li W.B., Becker K.G.,
RA Derby M.L., Ross C.A., Margolis R.L.;
RT "CCG repeats in cDNAs from human brain."
RL Hum. Genet. 103:666-673(1998).
RN [2]
RP CONCEPTUAL TRANSLATION OF 1-87.
RA Pardo B.;
RN Unpublished observations (JUN-2001).
CC -1- FUNCTION: NEURONAL CELL SURFACE PROTEIN THAT MAY BE INVOLVED IN
CC CELL RECOGNITION AND CELL ADHESION BY FORMING INTRACELLULAR
CC JUNCTIONS THROUGH BINDING TO NEUROLIGINS. MAY PLAY A ROLE IN
CC FORMATION OR MAINTENANCE OF SYNAPTIC JUNCTIONS. MAY MEDIATE
CC INTRACELLULAR SIGNALING.
CC -1- SUBUNIT: THE CYTOPLASMIC C-TERMINAL REGION BINDS TO CASR. SPECIFIC
CC ISOFORMS BIND NEUROLIGIN, ALPHA-DYSTROGLYCAN AND ALPHA-LATROTOXIN
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS: MULTIPLE ISOFORMS MAY BE PRODUCED BY
CC ALTERNATIVE SPLICING (BY SIMILARITY). BY USE OF ALTERNATIVE
CC PROMOTERS IN THE NEUREXIN 1 GENE, THE ALPHA (AC 09ULB1) AND BETA
CC ISOFORMS DIFFER IN THEIR N-TERMINI.
CC -1- PTM: HIGHLY O-GLYCOSYLATED AND MINOR N-GLYCOSYLATED (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NEUREXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```


fragment.";
 Proc. Natl. Acad. Sci. U.S.A. 94:10547-10553(1997).
 [4]
 COVALENT BOND.
 MEDLINE:99268331; PubMed:10338009;
 RA Buse G., Soulimane T., Demor M., Meyer H.E., Blueggel M.;
 RT "Evidence for a copper-coordinated histidine-tyrosine cross-link in
 RT the active site of cytochrome oxidase.";
 RT Protein Sci. 8:985-990(1999).
 [5]
 REVIEW.
 RA MEDLINE:96391941; PubMed:8794157;
 RA Ostermeier C., Iwata S., Michel H.;
 RT "Cytochrome c oxidase.";
 RT Curr. Opin. Struct. Biol. 6:460-466(1996).
 CC -I- FUNCTION: SUBUNIT I AND II FORM THE FUNCTIONAL CORE OF THE ENZYME
 CC COMPLEX. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED VIA
 CC HEME A AND CU(II) TO THE BINUCLEAR CENTER FORMED BY HEME A3 AND
 CC CU(II). THIS CYTOCHROME C OXIDASE SHOWS PROTON PUMP ACTIVITY
 CC ACROSS THE MEMBRANE IN ADDITION TO THE ELECTRON TRANSFER.
 CC -I- CATALYTIC ACTIVITY: 4 ferriocytochrome c + O(2) = 4 ferriocytochrome
 CC c + 2 H(2)O.
 CC -I- COFACTOR: TWO HEME GROUPS AND COPPER B.
 CC -I- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -I- PTR: HIS-276 AND TYR-280 RESIDUES ARE INVOLVED IN THE FORMATION OF
 CC A COPPER-COORDINATED COVALENT CROSSLINK AT THE ACTIVE SITE OF THE
 CC CATALYTIC SUBUNIT I.
 CC -I- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y07533; CAA68821.1; -.
 CC PIR: S08270; S08270.
 CC PDB: 1AR1; 1I-FEB-98.
 CC InterPro: IPR000883; COX1.
 CC Pfam: PF00115; COX1; 1.
 CC PRINTS: PR01165; CYCOXIDASE1.
 CC PROSITE: PS00077; COX1; 1.
 CC Oxioreductase; Heme; Copper; Ion transport; 3D-structure.
 KW Respiratory chain; Hydrogen
 FT DOMAIN 1 28
 FT TRANSMEM 29 59
 FT TRANSMEM 60 82
 FT TRANSMEM 83 120
 FT TRANSMEM 121 126
 FT TRANSMEM 127 131
 FT TRANSMEM 152 176
 FT TRANSMEM 177 206
 FT TRANSMEM 207 217
 FT TRANSMEM 218 251
 FT TRANSMEM 252 262
 FT TRANSMEM 263 299
 FT TRANSMEM 300 303
 FT TRANSMEM 304 331
 FT TRANSMEM 332 332
 FT TRANSMEM 333 364
 FT TRANSMEM 365 369
 FT TRANSMEM 370 395
 FT TRANSMEM 396 404
 FT TRANSMEM 405 437
 FT TRANSMEM 438 440
 FT TRANSMEM 441 469
 FT TRANSMEM 470 478
 FT TRANSMEM 479 514
 FT TRANSMEM 515 558
 FT METAL 94 94

FT	METAL	276	276	COPPER B.
FT	METAL	280	280	COPPER B.
FT	METAL	325	325	COPPER B.
FT	METAL	326	326	COPPER B.
FT	METAL	411	411	IRON (HEME A3).
FT	METAL	413	413	IRON (HEME A).
FT	DISULFID	66	80	
SO	SEQUENCE	558 AA;	62439 MW;	A8402453C0C0339E CRC64;

Query Match	Best Local Similarity	Score 31;	DB 1;	Length 558;
Matches 6;	Conservative	75.0%;	Pred. No. 28;	
		0;	Mismatches	2; Indels
				Gaps 0;

OY	1	YTAGOOLT 8		
DB	328	YTAGMSLT 335		

RESULT 15
 CH11_BACCI STANDARD; PRT; 699 AA.
 ID CH11_BACCI
 AC P20533;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Chitinase A1 precursor (EC 3.2.1.14).
 GN CH11.
 OS Bacillus circulans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID:1397;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WL-12.
 RA MEDLINE:90368776; PubMed:2203782;
 RA Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;
 RT "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed
 RT its evolutionary relationship to Serratia chitinase and to the type
 RT III homology units of fibronectin.";
 RT J. Biol. Chem. 265:15659-15665(1990).
 RN [2]
 RP MUTAGENESIS.
 RC STRAIN-WL-12.
 RA MEDLINE:93366760; PubMed:8103047;
 RA Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.,
 RA Uchida M., Tanaka H.;
 RT "Identification of glutamic acid 204 and aspartic acid 200 in
 RT chitinase A1 of Bacillus circulans WL-12 as essential residues for
 RT chitinase activity.";
 RT J. Biol. Chem. 268:18567-18572(1993).
 CC -I- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
 CC acetyl-D-glucosamine polymers of chitin.
 CC -I- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -I- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M57601; AAA81528.1; -.
 CC PIR: A38368; A38368.
 CC HSSP: P07254; 1CPN.
 CC InterPro: IPR003610; Chitin_bing3.
 CC InterPro: IPR001579; Chitinase_2.
 CC InterPro: IPR003961; FN_III.
 CC InterPro: IPR003962; FNIII_repeat.
 CC InterPro: IPR001223; Glyco_hydro_18.

DR Pfam; PF02839; CBD_7; 1.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00014; FNYPEIII.
DR SMART; SM00495; ChEBD3; 1.
DR SMART; SM00060; FN3; 2.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Repeat.
FT SIGNAL 1 41
FT CHAIN 42 699
FT DOMAIN 42 460
FT DOMAIN 465 549
FT DOMAIN 560 644
FT ACT_SITE 204 204
FT MUTAGEN 200 200
FT MUTAGEN 200 200
FT MUTAGEN 204 204
SQ SEQUENCE 699 AA; 73677 MW; ACTG9B22E2987643 CRC64;
CATALYTIC.
FIBRONECTIN TYPE-III (R-1).
FIBRONECTIN TYPE-III (R-2).
PROTON DONOR (PROBABLE).
D->N: DECREASE IN ACTIVITY.
D->E: NO CHANGE IN ACTIVITY.
E->D,Q: LOSS OF ACTIVITY.

Query Match

75.6%; Score 31; DB 1; Length 699;

Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGQLT 8
|||||:1
Db 662 YTAGQLVT 669

Search completed: June 13, 2002, 09:11:55
Job time: 1035 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 13, 2002, 08:55:55 ; Search time 73.31 Seconds
(without alignments)
10.486 Million cell updates/sec

Title: US-09-710-339-2_COPY_448_455
Perfect score: 41
Sequence: 1 YRAGQOLR 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	478	2 JK0201	alpha-amylose (EC
2	41	100.0	498	2 A48305	alpha-amylose (EC
3	41	100.0	499	1 ALAS1	alpha-amylose (EC
4	41	100.0	499	1 ALAS3	alpha-amylose (EC
5	41	100.0	499	2 B48305	alpha-amylose (EC
6	41	100.0	499	2 US0663	alpha-amylose (EC
7	41	100.0	499	2 JT0466	alpha-amylose (EC
8	41	100.0	499	2 JN0588	alpha-amylose (EC
9	34	82.9	849	2 D82510	chitinase VCA0027
10	33	80.5	384	2 H64161	hypothetical prote
11	32	78.0	390	2 T30395	probable transcrip
12	32	78.0	399	2 T41786	P47 orf40 - Bombyx
13	32	78.0	399	2 T10314	viral transcriptio
14	32	78.0	401	2 H72854	viral transcriptio
15	32	78.0	1250	2 T27706	hypothetical prote
16	31	75.6	214	2 D46681	hypothetical prote
17	31	75.6	214	2 E46681	hypothetical prote
18	31	75.6	221	2 F81876	hypothetical prote
19	31	75.6	242	2 T13190	hypothetical prote
20	31	75.6	282	2 H82394	hypothetical prote
21	31	75.6	347	2 D83653	oligopeptide ABC t
22	31	75.6	554	2 S03809	cytochrome-c oxida
23	31	75.6	558	2 S08270	cytochrome-c oxida
24	31	75.6	699	2 A38368	chitinase (EC 3.2.
25	30	73.2	3020	2 A49332	mucin 2 precursor,
26	30	73.2	104	2 T46246	hypothetical prote
27	30	73.2	108	2 E72766	hypothetical prote
28	30	73.2	255	2 D57143	penicillin-binding
29	30	73.2	283	2 F87323	hypothetical prote

30	30	73.2	289	2 A10446	probable carbon-ni
31	30	73.2	331	2 H83240	probable ATP-bindi
32	30	73.2	339	2 AF0300	probable oxidoredu
33	30	73.2	388	2 A75318	conserved hypothet
34	30	73.2	410	2 T36610	hypothetical prote
35	30	73.2	425	2 E71075	hypothetical prote
36	30	73.2	449	2 A24993	probable proton g1
37	30	73.2	452	2 H84772	cellulase (EC 3.2.
38	30	73.2	474	2 S41117	probable serine ca
39	30	73.2	505	2 AC1469	nitric-oxide reduc
40	30	73.2	507	2 T35677	intercalin like pr
41	30	73.2	609	2 D87049	probable hydrolase
42	29	70.7	53	2 S78004	hypothetical prote
43	29	70.7	160	2 F70049	cuticle structural
44	29	70.7	191	2 D82425	flagellin homolog
45	29	70.7	211	2 E83116	guanylate cyclase- 50S ribosomal prot

ALIGNMENTS

RESULT 1

JK0201
alpha-amylose (EC 3.2.1.1) - Aspergillus oryzae

N:Alternate names: glycoenzyme; Taka-amylose A
C:Species: Aspergillus oryzae

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JK0201

R:Toda, H.; Kondo, K.; Narita, K.
Proc. Jpn. Acad. 58B, 208-212, 1982

A:Title: The complete amino acid sequence of Taka-amylose A.
A:Reference number: JK0201

A:Accession: JK0201
A:Molecule type: protein

A:Residues: 1-478 <TOD>
C:Comment: One atom of calcium per molecule is essential for the activity.

C:Comment: This enzyme is a glycoprotein.
C:Comment: See also PIR:JT0466 and PIR:JS0240.

C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:173-300/Domains: alpha-amylose core homology <AMY>

F:197/Binding site: carbohydrate (asn) (covalent) #status experimental
F:210/230/257/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 41; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAGQOLR 8
DB 427 YRAGQOLR 434

RESULT 2
A48305

alpha-amylose (EC 3.2.1.1) A precursor - Aspergillus awamori
C:Species: Aspergillus awamori

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: A48305

R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.
Curr. Genet. 17, 203-212, 1990

A:Title: Cloning, characterization, and expression of two alpha-amylose genes from As
A:Reference number: A48305; MUID:90254827

A:Accession: A48305
A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA
A:Residues: 1-498 <KOR>

C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 41; DB 2; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
 |||||
 Db 449 YTAGOQLT 456

RESULT 3
 ALAS1
 alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2000
 C:Accession: S04548; A33214; J50240; A91930; A93767; A10627

R:Witsel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon
 A:Reference number: S04548; MUID:89378767

A:Accession: S04548
 A:Molecule type: DNA

A:Residues: 1-499 <WR>
 A:Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA1218.1; PID:g295921

A:Genetics: AMY1

A:Accession: A33214
 A:Molecule type: mRNA

A:Residues: 1-499 <W12>
 A:Cross-references: GB:X12725; NID:g2430; PIDN:CAA1218.1; PID:g295921

R:Gene, M.J.; Dove, M.J.; Seligy, V.L.
 Gene 79, 107-117, 1989

A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
 A:Reference number: J50240; MUID:89378767

A:Accession: J50240
 A:Molecule type: DNA

A:Residues: 1-499 <GEN>
 A:Genetics: AMY2

A>Note: the authors refer to this as isozyme II

R:Isemura, S.; Ikenaka, T.
 J. Biochem. 74, 1-10, 1973

A:Reference number: A91930; MUID:74001521
 A:Accession: A91930

A:Molecule type: protein
 A:Residues: 206-225 <ISE>

R:Marita, K.
 Proc. Jpn. Acad. 51, 285-290, 1975

A:Reference number: A93767
 A:Accession: A93767

A:Molecule type: protein
 A:Residues: 434-443, 446-447, 'Q', 449-458, 'GTRV', 459-464, 467-468, 'B', 470, 'B', 472-499 <NAR>

J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylase A.

A:Reference number: A37454; MUID:84212370
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms

R:Swift, H.J.; Brady, L.; Derenda, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.
 submitted to the Brookhaven Protein Data Bank, August 1992

A:Reference number: A51548; PDB:6FRA
 A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497

C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: <AMY>

A:Gene: amy1
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Genetics: <AMY>
 A:Gene: amy2; Amy1
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po

F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase 1 #status experimental <MAT>

F:194-321/Domain: alpha-amylase core homology <AMY>
 F:51-59,171-185,261-304,461-496/disulfide bonds: #status experimental

F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 41; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
 |||||
 Db 449 YTAGOQLT 456

RESULT 4
 ALAS3
 alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713

R:Witsel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
 A:Reference number: S04548; MUID:89378767

A:Accession: S04549
 A:Molecule type: DNA

A:Residues: 1-499 <WR>
 A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA1220.1; PID:g295922

A:Accession: A33215
 A:Status: translation not shown

A:Molecule type: mRNA
 A:Residues: 1-499 <W12>

A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA1220.1; PID:g295922
 R:Gene, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each contain

A:Reference number: J50240; MUID:89378767
 A:Accession: A44713

A:Molecule type: DNA
 A:Residues: 1-499 <GEN>

A>Note: the authors refer to this as isozyme I
 R:Matsura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.

J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylase A.

A:Reference number: A37454; MUID:84212370
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms

C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: amy3

A:Gene: amy3; Amy1
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
 F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase 3 #status experimental <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>

F:51-59,171-185,261-304,461-496/disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted

F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 41; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||||
DB 449 YTAGOQLT 456

RESULT 5
B48305
alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori

C:Species: Aspergillus awamori
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: B48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus
A:Reference number: A48305; MUID:90234827
A:Accession: B48305
A:Molecule type: DNA
A:Status: preliminary; not compared with conceptual translation
A:Residues: 1-499 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
F:194-321/Product: alpha-amylase #status predicted <AMY>
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 41; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||||
DB 449 YTAGOQLT 456

RESULT 6

JS0663
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
C:Species: Aspergillus sp.
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
C:Accession: JS0663
R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Biotech. Biotechnol. Biochem. 56, 174-179, 1992
A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousamii and its expression
A:Reference number: JS0663; MUID:9232146
A:Accession: JS0663
A:Molecule type: mRNA
A:Residues: 1-499 <SHI>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <ALP>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 41; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||||
DB 449 YTAGOQLT 456

RESULT 7
JT0466
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae

N:Alternate names: glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JT0466
R:Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus
A:Reference number: JT0466
A:Accession: JT0466
A:Molecule type: DNA
A:Residues: 1-499 <TAD>
C:Comment: See also PIR:JK0201 and PIR:JS0240.
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 41; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||||
DB 449 YTAGOQLT 456

RESULT 8

JN0588
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
C:Accession: JN0588
R:Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsudoi, A.; Udaoka, S.
Gene 84, 319-327, 1989
A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
A:Reference number: JN0588; MUID:90128276
A:Accession: JN0588
A:Molecule type: mRNA
A:Residues: 1-499 <TSU>
C:Comment: The alpha amylases are encoded by multigene family.
C:Genetics:
A:Gene: Taa-G1
A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 41; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8

Db 449 YTAGOQLT 456
|||||
RESULT 9
D82510
Chinese VCA0027 [Imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82510
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, H.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: D82510
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-849 <HEI>
A:Cross-references: GB:AE004346; GB:AE003853; NID:96657401; PIDN:AAF95941.1; GSPDB:GN001
C:Genetics:
A:Gene: VCA0027
A:Map position: 2
Query Match 82.9%; Score 34; DB 2; Length 849;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;
Indels 0;
OY 1 YTAGOQLT 8
|||||
Db 805 YTGGOQVT 812
RESULT 10
H64161
Hypothetical protein HI0934 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
C:Accession: H64161
R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Cocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: H64161
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-384 <TIGR>
A:Cross-references: GB:U32775; GB:U42023; NID:91573951; PIDN:AAC22592.1; PID:91573955; T
Query Match 80.5%; Score 33; DB 2; Length 384;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGOQLT 8
|||||
Db 301 YAGGOQMT 308
RESULT 11
T30395
probable transcription regulator p47 - Lymantria dispar nuclear polyhedrosis virus
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30395
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavecek, J.M.; Rohl
Virology 253, 17-34, 1999

A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria
A:Reference number: Z20836; MUID:99124785
A:Accession: T30395
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-390 <KU2>
A:Cross-references: EMBL:AF081810; NID:93822234; PIDN:AAC70233.1; PID:93822282
C:Keywords: transcription factor
Query Match 78.0%; Score 32; DB 2; Length 390;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGOQLT 8
|||||
Db 140 YTLGOQLS 147
RESULT 12
T41786
P47 orf40 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BMSNPV
A:Variety: Isolate T3
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T41786
R:Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: Z22020; MUID:99281911
A:Accession: T41786
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-399 <KAM>
A:Cross-references: EMBL:L33180; NID:93745835; PIDN:AAC63715.1; PID:93745868
A:Experimental source: Isolate T3
C:Genetics:
A:Note: p47
Query Match 78.0%; Score 32; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGOQLT 8
|||||
Db 144 YTLGOQLS 151
RESULT 13
T10314
viral transcription regulator p47 - Orgyia pseudotsugata nuclear polyhedrosis virus
C:Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OPMNPV
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10314
R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F.
Virology 229, 381-399, 1997
A:Title: The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedro
A:Reference number: Z17011; MUID:97271300
A:Accession: T10314
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-399 <AHR>
A:Cross-references: EMBL:U75930; NID:92934903; PIDN:AAC59044.1; PID:91911291
C:Keywords: transcription regulation
Query Match 78.0%; Score 32; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGOQLT 8
|||||

Db 148 YTLGQQLS 155

RESULT 14

H72854

Viral transcription regulator p47 - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

A:Note: dsDNA virus

C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 21-Jul-2000

C:Accession: H72854; B45684

R:Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.

Virology 202, 586-605, 1994

A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.

A:Reference number: A72850; MUID:94303173

A:Accession: H72854

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-401 <AVR>

A:Cross-references: GB:L22858; NID:g510708; PIDN:AAA66670.1; PID:g559109

R:Garstens, E.B.; Lu, A.L.; Chan, H.L.

J. Virol. 67, 2513-2520, 1993

A:Title: Sequence, transcriptional mapping, and overexpression of p47, a baculovirus gen

A:Reference number: A45684; MUID:93233214

A:Accession: B45684

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-401 <CAK>

A:Cross-references: GB:L07878; NID:g293990; PIDN:AAA16859.1; PID:g293992

A:Note: sequence extracted from NCBI backbone (NCBIN:129587, NCBIF:129589)

C:Genetics:

A:Gene: Ac-p47

C:Keywords: nucleus; transcription regulation

Query Match 78.0%; Score 32; DB 2; Length 401;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 15

T27706

hypothetical protein ZK1193.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27706

R:Geisler, C.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid ZK1193.

A:Reference number: Z20409

A:Accession: T27706

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1250 <GEI>

A:Cross-references: EMBL:U41553; PIDN:AAA83291.1; CESP:ZK1193.2

C:Genetics:

A:Gene: CESP:ZK1193.2

A:introns: 95/2; 228/1; 302/1; 355/2; 388/1; 400/1; 451/1; 514/1; 593/1; 643/3; 822/2; 8

Query Match 78.0%; Score 32; DB 2; Length 1250;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGQO 6
|||||
Db 698 YTAGQO 703

Fri Jun 14 10:36:50 2002

us-09-710-339-2_copy_448_455.rpr

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:53:29 ; Search time 158.68 seconds
(without alignments)
5.600 Million cell updates/sec

Title: us-09-710-339-2_COPY_448_455

Perfect score: 41

Sequence: 1 YTAGQQLR 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: A_Geneseq.032802:*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	478	15	AA1980.DAT
2	41	100.0	478	16	AA1981.DAT
3	41	100.0	478	16	AA1982.DAT
4	41	100.0	478	16	AA1983.DAT
5	41	100.0	478	17	AA1984.DAT
6	41	100.0	498	22	AA1985.DAT
7	32	78.0	1626	22	AA1986.DAT
8	31	75.6	52	22	AA1987.DAT
9	31	75.6	215	22	AA1988.DAT
10	31	75.6	215	22	AA1989.DAT
11	31	75.6	223	22	AA1990.DAT

12	31	75.6	302	22	AA1991.DAT
13	31	75.6	342	16	AA1992.DAT
14	31	75.6	343	17	AA1993.DAT
15	31	75.6	746	22	AA1994.DAT
16	31	75.6	746	22	AA1995.DAT
17	31	75.6	746	22	AA1996.DAT
18	31	75.6	858	22	AA1997.DAT
19	31	75.6	1327	22	AA1998.DAT
20	31	75.6	1522	22	AA1999.DAT
21	31	75.6	5179	22	AA2000.DAT
22	31	75.6	58	22	AA2001.DAT
23	30	73.2	59	22	AA2002.DAT
24	30	73.2	98	21	AA2003.DAT
25	30	73.2	109	18	AA2004.DAT
26	30	73.2	222	22	AA2005.DAT
27	30	73.2	222	22	AA2006.DAT
28	30	73.2	222	22	AA2007.DAT
29	30	73.2	279	22	AA2008.DAT
30	30	73.2	331	22	AA2009.DAT
31	30	73.2	449	18	AA2010.DAT
32	30	73.2	452	21	AA2011.DAT
33	30	73.2	456	21	AA2012.DAT
34	30	73.2	466	22	AA2013.DAT
35	30	73.2	472	22	AA2014.DAT
36	30	73.2	472	22	AA2015.DAT
37	30	73.2	630	18	AA2016.DAT
38	30	73.2	630	18	AA2017.DAT
39	30	73.2	653	19	AA2018.DAT
40	30	73.2	2237	22	AA2019.DAT
41	29	70.7	91	22	AA2020.DAT
42	29	70.7	102	22	AA2021.DAT
43	29	70.7	156	21	AA2022.DAT
44	29	70.7	156	21	AA2023.DAT
45	29	70.7	156	21	AA2024.DAT

ALIGNMENTS

RESULT 1
ID AAR46065 standard; protein; 478 AA.
AC AAR46065;
DT 18-JUL-1994 (first entry)
XX
DE Mutant alpha-amylase.
KW Methionine substitution; stability; activity; detergent;
KW dishwashing agents; liguifaction agents.
XX
OS Aspergillus oryzae.
XX
FN WO9402597-A.
XX
PD 03-FEB-1994.
XX
PE 06-JUL-1993; 93WO-DK00230.
XX
PR 23-JUL-1992; 92DK-0000946.
PR 16-DEC-1992; 92DK-0001503.
PR 15-MAR-1993; 93DK-0000292.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Bisgard-frantzen H, Svendsen A;
XX
XX WPI; 1994-048855/06.
XX
XX Mutant alpha-amylase from Bacillus species comprising a
XX methionine substitution - with improved stability and activity at
XX low pH, for use in detergents, dishwashing agents and

PT liquefaction agents
XX
PS Claim 1; Page 7; 20pp; English.
XX
CC The sequence as that of the *Aspergillus oryzae* alpha amylase, sold
CC commercially as FRINGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC be mutated by substitution of one or more of its methionine residues
CC for any amino acid other than cysteine. The mutant alpha-amylase
CC exhibits a better activity level and better stability in the
CC presence of oxidizing agents than previous mutant alpha amylases,
CC and improved thermostability at moderately low pH. The enzyme can
CC be used as an additive for detergents, dishwashing agents and
CC liquefaction agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 41; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

OY 1 YTAGOQLT 8
Db 428 ytaagqlt 435
|||||||

RESULT 2
AAR79025
ID AAR79025 standard; protein; 478 AA.
XX
AC AAR79025;
XX
DT 22-MAR-1996 (first entry)
XX
DE Mature taka-amylase A.
XX
KW Wild type: neopullulanase; B. stearotherophilus; mutant;
KW food industry; modification; hydrophobicity; replacement
KW insertion; deletion.
XX
OS *Aspergillus oryzae*.
XX
FH Key Location/Qualifiers
FH Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
PN JP07177891-A.
XX
PD 18-JUL-1995.
XX
PF 31-OCT-1994; 94JP-0288658.
XX
PR 12-NOV-1993; 93JP-0306096.
XX
PA (EZA) EZA KI GLICO CO.
PA (NTDE) NEC CORP.
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site - increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
PS Disclosure; Page 10-11; 18pp; English.
XX
CC This sequence represents the mature form of taka-amylase from *A. oryzae*.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the

CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC hydrophobic amino acid from the selected site. The method was used
CC to produce neopullulanases Y377F, S422V and M375L.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 41; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

OY 1 YTAGOQLT 8
Db 427 ytaagqlt 434
|||||||

RESULT 3
AAR72450
ID AAR72450 standard; Protein; 478 AA.
XX
AC AAR72450;
XX
DT 01-DEC-1995 (first entry)
XX
DE *Aspergillus oryzae* alpha amylase (mature protein).
XX
KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
KW Bacillus stearotherophilus; dyeing; bleaching; scouring; textile;
KW thermostable.
XX
OS *Aspergillus oryzae*.
XX
PN WO9510603-A.
XX
PD 20-APR-1995.
XX
PF 05-OCT-1994; 94WO-DK00370.
XX
PR 08-OCT-1993; 93DK-0001133.
PR 02-FEB-1994; 94DK-0000140.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Svendsen A, Thellersen M;
PI Van der zee P, Bisgardfrantzen H, Borchert T;
XX
DR WPI; 1995-161790/21.
XX
PT New Bacillus derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance
XX
PS Disclosure; Page 75-76; 105pp; English.
XX
CC Variant alpha amylase enzymes which have improved washing and/or
CC as detergent additives. The enzymes have one or more amino acid
CC residues added, deleted or substituted. The variants can also be
CC used for textile desizing prior to scouring, bleaching and dyeing.
CC The variants have improved thermostability, acid/alkaline stability;
CC low temperature optimum; pH optimum; higher hydrolysis velocity and
CC improved tolerance to other composition constituents, e.g. oxidation
CC agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 41; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

OY 1 YTAGOQLT 8
|||||||

Db 428 ytagqqlt 435

RESULT 4
 AAR78270
 ID AAR78270 standard; protein; 478 AA.
 XX
 AC AAR78270;
 XX
 DT 17-JAN-1996 (first entry)
 XX
 DE Aspergillus oryzae alpha amylase (mature protein).
 XX
 KM Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
 KM starch; thermostable; methionine; Bacillus licheniformis;
 KM Bacillus amyloliquefaciens; Bacillus steaerothermophilus;
 KM Aspergillus oryzae.
 XX
 OS Aspergillus oryzae.
 XX
 FN W09521247-A1.
 XX
 PD 10-AUG-1995.
 XX
 PF 05-OCT-1994; 94WO-DK00371.
 XX
 PR 02-FEB-1994; 94DK-0000141.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Marcher D, Nilsson TE, Pedersen HH, Toft AH;
 DR WPI: 1995-263767/37.
 XX
 PT Use of an oxidation stable alpha-amylase - for simultaneous desizing
 PT and bleaching or scouring of fabrics contg. starch or starch derivs.
 XX
 PS Disclosure; Page 25-26; 37pp; English.
 XX
 SS Oxidation stable alpha amylases can be used for the simultaneous
 CC desizing and bleaching or scouring of a fabric comprising starch or
 CC starch derivatives. They exhibit a better heat stability
 CC especially in the presence of oxidizing agents. They are obtained
 CC from a parent alpha amylase by replacing one or more methionine
 CC residues with any amino acid different from Cys or Met, preferably
 CC Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is
 CC pref. derived from a Bacillus species, although alpha amylases of
 CC fungal origin can also be used. This sequence is the wild type
 CC (unmodified) alpha amylase of Aspergillus oryzae.
 CC
 SQ Sequence 478 AA;
 SO

Query Match 100.0%; Score 41; DB 16; Length 478;
 Best Local Similarity 100.0%; Pred. No. 0.78;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
 |||||
 Db 428 ytagqqlt 435

RESULT 5
 AAW14500
 ID AAW14500 standard; protein; 478 AA.
 XX
 AC AAW14500;
 XX
 DT 04-JUN-1997 (first entry)
 XX
 DE Aspergillus oryzae alpha-amylase (mature protein).
 XX
 KM alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;

KW Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;
 KW calcium dependency; substrate binding; stability; pH optimum;
 KW thermostability; cleavage; oligosaccharide substrate; dishwashing;
 KW washing; detergent additive; fabric desizing; starch liquefaction;
 KW sweetener; ethanol production; variant.
 XX
 OS Aspergillus oryzae.
 XX
 FN Key
 XX Location/Qualifiers
 FH
 FT MISC-difference 13..45
 FT /label= loop 1 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 7-23 of AAW14499 is deleted or
 FT replaced with a fragment corresponding to
 FT this fragment; claim 33"
 FT
 FT MISC-difference 14..40
 FT /label= loop 1 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 8-18
 FT of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim
 FT 35"
 FT MISC-difference 28..42
 FT /label= loop 1 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 12-19 of AAW14499 is deleted or
 FT replaced with a fragment corresponding to
 FT this fragment; claim 30"
 FT
 FT MISC-difference 32..38
 FT /label= loop 1 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 14-15
 FT of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim
 FT 32"
 FT MISC-difference 66..84
 FT /label= loop 2 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 44-57 of AAW14499 is deleted or
 FT replaced with a fragment corresponding to
 FT this fragment; claim 18"
 FT
 FT MISC-difference 70..78
 FT /label= loop 2 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 48-51
 FT of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim
 FT 20"
 FT MISC-difference 98..210
 FT /label= loop 3 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 117-185 of AAW14499 is deleted
 FT or replaced with a fragment corresponding to
 FT this fragment; claim 24"
 FT
 FT MISC-difference 102..206
 FT /label= loop 3 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 121-181
 FT of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim
 FT 26"
 FT MISC-difference 121..181
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to this fragment is deleted or
 FT replaced with a fragment corresponding to
 FT 102-206 of AAW14499; claim 41"
 FT
 FT MISC-difference 121..174
 FT /note= "preferred region where at least one amino acid

```

FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to this
FT fragment is deleted or replaced with a fragment
FT corresponding to 102-199 of AAM14499; claim 42"
FT
FT Misc-difference 165..177
FT /label= loop 3 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 195-202 of AAM14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 21"
FT
FT Misc-difference 166..173
FT /label= loop 3 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 196-198
FT of AAM14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 23"
FT Misc-difference 181..184
FT /note= "an amino acid fragment corresponding to this
FT region is deleted from the parent sequence of a
FT variant Fungamyl; claim 43"
FT
FT Misc-difference 291..313
FT /label= loop 8 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 322-346 of AAM14498 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 36"
FT
FT Misc-difference 297..313
FT /label= loop 8 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 325-345
FT of AAM14498 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 38"
FT
XX WO9623874-A1.
XX
XX 08-AUG-1996.
XX
XX
XX 05-FEB-1996; 96WO-DK00057.
XX
XX 10-NOV-1995; 95DK-0001256.
XX 03-FEB-1995; 95DK-0000128.
XX 23-OCT-1995; 95DK-0001192.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Bisgard-frantzen H, Borchert TV, Svendsen A;
XX WPI; 1996-371424/37.
XX
XX Alpha-amylase variants and methods of production - have altered
XX PT properties such as calcium dependency, substrate binding and
XX PT stability
XX
XX disclosure: Page 87-88; 171pp; English.
XX
XX The present sequence is the mature Aspergillus oryzae alpha-amylase (A).
XX CC Variants of parent termamyl- and fungamyl-like alpha-amylases (and
XX CC methods of constructing them) are claimed. Examples of variants are
XX CC featured above. The variants have altered properties such as calcium
XX CC dependency, substrate binding and stability. Also one or more proline or
XX CC cysteine residues in the variant is modified or replaced with a
XX CC non-proline or non-cystein residue such as alanine. The variants can be
XX CC used for (dish)washing, as detergent additives or for fabric desizing or
XX CC starch liquefaction. They can also be used for the production of
XX CC sweeteners and ethanol from starch. See also AAM14498-99.
XX
SQ Sequence 478 AA:

Query Match 100.0%; Score 41; DB 17; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGQQLT 8
DB 428 YTAGQQLT 435

RESULT 6
AAB84206
ID AAB84206 standard; Protein: 498 AA.
AC AAB84206;
XX
XX 06-AUG-2001 (first entry)
DT
XX
XX Amino acid sequence of a fungamyl-like alpha-amylase.
DE
XX Fungamyl-like alpha-amylase; glucoamylase; dextrinisation; maltose;
XX alcohol; starch; dough improver; brewing; starch liquefaction.
KW
XX Aspergillus oryzae.
XX
XX WO200134784-A1.
XX
XX 17-MAY-2001.
XX
XX 10-NOV-2000; 2000WO-DK00626.
XX
XX 10-NOV-1999; 99DK-0001617.
XX
XX (NOVO ) NOVOZYMES AS.
XX
XX Bisgard-Frantzen H, Svendsen A, Pedersen S;
XX WPI; 2001-367478/38.
XX
XX N-PSDB; AAF90208.
XX
XX New variant of Fungamyl-like alpha-amylase, useful for production of
XX maltose syrups, includes mutations that improve stability against heat
XX and acidic pH.
XX
XX Claim 1; Page 42-45; 49pp; English.
XX
XX The present sequence represents a fungamyl-like alpha-amylase. The
XX CC specification describes variants of this fungamyl-like alpha-amylase,
XX CC which have an alteration in one the amino acid regions 98-110, 150-160,
XX CC 161-167, 280-288, 448-455, and 468-473. Each alteration is a deletion
XX CC or substitution of an amino acid or an insertion of an amino acid
XX CC downstream of a particular position. The variants retain alpha-amylase
XX CC activity, and have better heat stability and/or stability at acidic pH,
XX CC relative to wild-type enzyme. The variants can therefore be used at
XX CC higher temperatures (more efficient conversion or faster reaction, and
XX CC have reduced need for cooling and reduced risk of contamination). The
XX CC variants may also be used in conjunction with other enzymes,
XX CC particularly glucoamylase during dextrinisation. The variants are
XX CC used to produce syrups, particularly of high maltose content, or alcohol,
XX CC from starch; as dough improver for baked goods; in brewing, to increase
XX CC fermentability of the wort; and for liquefaction of starch.
XX
SQ Sequence 498 AA:

Query Match 100.0%; Score 41; DB 22; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGQQLT 8
DB 448 YTAGQQLT 455

```

RESULT 7
 ABG20974
 ID ABG20974 standard; Protein; 1626 AA.
 XX
 AC ABG20974;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #20965.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 OS
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSED INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 P1
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS85161.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID NO 51333; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 1626 AA;

Query Match 78.0%; Score 32; DB 22; Length 1626;
 Best Local Similarity 85.7%; Pred. No. 3.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOQL 7
 |||||:
 DB 165 ytagqql 171

RESULT 8
 AAG64426
 ID AAG64426 standard; Peptide; 52 AA.
 XX
 AC AAG64426;
 XX
 DT 19-SEP-2001 (first entry)
 XX
 DE B. circulans chitin-cellulose binding domain (CBD) peptide.
 DE
 KW Heteromer peptide; chitin cellulose binding domain; CBD;
 KW protein immobilisation.
 XX
 OS Bacillus circulans.
 OS
 PN WO200131038-A1.
 XX
 PD 03-MAY-2001.
 XX
 PF 19-OCT-2000; 2000WO-JP07275.
 XX
 PR 22-OCT-1999; 99JP-0301699.
 XX
 PA (FUJII) FUJISAWA PHARM CO LTD.
 XX
 PI Tanaka A, Ueda M, Nagao K;
 P1
 DR WPI; 2001-417561/44.
 DR
 PT Heteromer peptides produced by auto-splicing of a precursor comprise
 PT subunits with chitin cellulose binding domain
 XX
 PS Disclosure; Page 85-86; 100pp; Japanese.
 XX
 CC The present sequence is that of the Bacillus circulans chitin-cellulose
 CC binding domain (CBD) peptide of the invention. The invention relates to
 CC heteromer peptides, produced by cleaving precursor peptides, comprise at
 CC least one subunit with an added chitin cellulose binding domain (CBD)
 CC useful for immobilising proteins for industrial use. The CBD fusion
 CC allows immobilisation of proteins without loss of activity and
 CC immobilised heteromer peptides may be produced with high yield.
 CC
 SQ Sequence 52 AA;

QY 1 YTAGOQL 8
 |||||:
 DB 15 ytagqlvt 22

Query Match 75.6%; Score 31; DB 22; Length 52;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 9
 ABB65782
 ID ABB65782 standard; Protein; 215 AA.
 XX
 AC ABB65782;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 24138.
 DE
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 PN
 PD 27-SEP-2001.
 PD
 XX

PF 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL09885.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 PS Disclosure: SEQ ID NO 24138; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB157737-AB12072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 215 AA:

Query Match 75.6%; Score 31; DB 22; Length 215;
 Best Local Similarity 75.0%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGGQQLT 8
 1 11 111
 DB 146 yvagsqtl 153

RESULT 10
 ABB71287
 ID ABB71287 standard; Protein: 215 AA.
 XX
 AC ABB71287;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 40653.
 DE
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL15390.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 PS Disclosure: SEQ ID NO 40653; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB157737-AB12072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 215 AA:

Query Match 75.6%; Score 31; DB 22; Length 215;
 Best Local Similarity 75.0%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGGQQLT 8
 1 11 111
 DB 146 yvagsqtl 153

RESULT 11
 ABB66336
 ID ABB66336 standard; Protein: 223 AA.
 XX
 AC ABB66336;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 23800.
 DE
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL10439.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 PS Disclosure: SEQ ID NO 25800; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 223 AA;

Query Match 75.6%; Score 31; DB 22; Length 223;
 Best Local Similarity 75.0%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGQOQLT 8
 |||||
 Db 151 yvagnqtl 158

RESULT 12

ABBS9066
 ID ABB59066 standard; Protein; 302 AA.
 XX

AC ABB59066;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 3990.

KW Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL03169.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure: SEQ ID NO 3990; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
 CC sequences (ABBS7737-ABB72072).
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC

SO Sequence 302 AA;

Query Match 75.6%; Score 31; DB 22; Length 302;
 Best Local Similarity 85.7%; Pred. No. 86;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGQOQL 7
 |||||
 Db 79 ylpqgql 85

RESULT 13

AAR66213
 ID AAR66213 standard; Protein; 342 AA.
 XX

AC AAR66213;

DT 04-AUG-1995 (first entry)

DE Nocardia corallina alkene monooxygenase subunit-1.

KW alkene monooxygenase; subunit-1; epoxidation; amoA gene.

OS Nocardia corallina.

PN JP06292571-A.

PD 21-OCT-1994.

PP 06-APR-1993; 93JP-0105171.

PR 06-APR-1993; 93JP-0105171.

PA (NIHA) JAPAN ENERGY CORP.

DR WPI; 1995-009069/02.

DR N-PSDB; AAQ79569.

PT Alkene monooxygenase and corresp. gene - useful for the
 PT epoxidation of an alkene

PS Claim 1; Page 2; 30pp; Japanese.

CC E.coli transformed with the DNA sequence AAQ79569 are able to
 CC catalyse the epoxidation of alkenes. The DNA is derived from
 CC Nocardia corallina and comprises 4 open reading frames. ORFs amoA
 CC and amoc encode subunits 1 and 2 of the alkene monooxygenase enzyme
 CC and ORF amod encodes a reductase capable of transferring electrons
 CC from NADH coenzyme to a monooxygenase.
 CC

SO Sequence 342 AA;

Query Match 75.6%; Score 31; DB 16; Length 342;
 Best Local Similarity 75.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGQOQLT 8
 |||||
 Db 34 ylvvgqst 41

RESULT 14

AAR81469
 ID AAR81469 standard; Protein; 343 AA.
 XX

AC AAR81469;

DT 07-AUG-1996 (first entry)

DE Nocardia corallina alkene mono-oxygenase gene product, amoA.

KW Alkene mono-oxygenase; indole; indigo production; biosynthesis;
 KW microbial oxidation; dye.

OS Nocardia corallina B-276.

PN JP08023988-A.

XX 30-JAN-1996.
PD
XX
XX 08-JUL-1994; 94JP-0179688.
PF
XX
XX 08-JUL-1994; 94JP-0179688.
PR
XX
XX (NITHA) JAPAN ENERGY CORP.
PA
XX WPI; 1996-133426/14.
DR
XX N-PSDB; AAT17418.
DR
XX
XX Prep'n. of indigo by a microbiological method - by culturing a
PT microbe having alkene monooxygenase activity to oxidise indole to
PT indigo
PS
XX Claim 5; Page 6-7; 11pp; Japanese.
PS
XX AAR81469-R81472 are protein products of the Nocardia corallina strain
CC B-276 alkene mono-oxygenase gene. The gene encodes 4 protein products
CC amca, amob, amoc and amod derived from the 3 different reading frames
CC of the operon. The gene is useful for the production of indigo via
CC oxidation of indole. Nocardia corallina can be cultured in a medium
CC contg. indole and will readily oxidise the indole yielding indigo into
CC the culture medium. E. coli may also be transformed with the alkene
CC mono-oxygenase gene and used as above to efficiently produce indigo
CC by microbial oxidation.
XX
XX
SQ Sequence 343 AA;

Query Match 75.6%; Score 31; DB 17; Length 343;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 YTAGQQLT 8
||| |
DB 35 ylvqqst 42

RESULT 15
AAG64423
XX AAG64423 standard; Protein; 746 AA.
XX
XX AAG64423;
AC
XX
XX 19-SEP-2001 (first entry)
DT
XX
XX Chimeric CBD-fused GL-7ACA acylase 1.
DE
XX
XX Heteromer peptide; chitin cellulose binding domain; CBD;
KW protein immobilisation.
XX
XX Chimeric - Bacillus circulans.
OS
XX Chimeric - Unidentified.
OS
XX WO200131038-A1.
PN
XX
XX 03-MAY-2001.
PD
XX
XX 19-OCT-2000; 2000WO-JP07275.
PF
XX
XX 22-OCT-1999; 99JP-0301699.
PR
XX
XX (FUJIT) FUJISAMA PHARM CO LTD.
PA
XX
XX Tanaka A, Ueda M, Nagao K;
PI
XX
XX WPI; 2001-417561/44.
DR
XX N-PSDB; AAH46095.
XX
XX Heteromer peptides produced by auto-splicing of a precursor comprise
PT subunits with chitin cellulose binding domain -

XX Claim 8; Page 68-71; 100pp; Japanese.
PS
XX
XX The present sequence is that of the CBD-fused GL-7ACA acylase chimeric
CC protein of the invention. The invention relates to heteromer peptides;
CC produced by cleaving precursor peptides, comprise at least one subunit
CC with an added chitin cellulose binding domain (CBD) useful for
CC immobilising proteins for industrial use. The CBD fusion allows
CC immobilisation of proteins without loss of activity and immobilised
CC heteromer peptides may be produced with high yield.
XX
XX
SQ Sequence 746 AA;

Query Match 75.6%; Score 31; DB 22; Length 746;
Best Local Similarity 75.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 YTAGQQLT 8
||| | : |
DB 16 ytagqlvt 23

Search completed: June 13, 2002, 08:53:30
Job time: 280 sec

Fri Jun 14 10:36:48 2002

us-09-710-339-2_copy_448_455.rag

Page 9

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:29 ; Search time 57.95 Seconds
(without alignments)
5.479 Million cell updates/sec

Title: US-09-710-339-2_COPY_98_110
Perfect score: 78
Sequence: 1 AYHGWOODIYSL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCUTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/bcckfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	78	100.0	478	1	US-08-720-899-7
2	78	100.0	478	1	US-08-459-610-7
3	78	100.0	478	2	US-08-343-804-7
4	78	100.0	478	2	US-08-600-808A-10
5	78	100.0	478	3	US-08-683-838A-10
6	78	100.0	478	4	US-08-182-859-7
7	62	79.5	478	2	US-08-339-715A-2
8	54	69.2	468	1	US-08-204-656B-2
9	54	69.2	468	1	US-08-470-702-6
10	54	69.2	468	1	US-08-467-831-6
11	53	67.9	468	1	US-08-204-656B-4
12	53	67.9	468	1	US-08-470-702-7
13	53	67.9	468	1	US-08-467-831-7
14	50	64.1	468	1	US-08-204-656B-6
15	50	64.1	468	1	US-08-470-702-8
16	49	62.8	468	1	US-08-467-831-8
17	49	62.8	468	1	US-08-204-656B-8
18	49	62.8	468	1	US-08-470-702-9
19	49	62.8	468	1	US-08-467-831-9
20	49	62.8	685	3	US-08-947-965-72
21	49	62.8	685	3	US-08-947-965-74
22	49	62.8	686	3	US-08-947-965-70
23	49	62.8	686	3	US-08-947-965-73
24	47	60.3	624	3	US-08-947-965-78
25	47	60.3	655	1	US-08-469-202-27
26	47	60.3	655	1	US-08-469-202-28
27	47	60.3	655	2	US-08-484-434C-34

28	47	60.3	655	2	US-08-484-434C-35	Sequence 35, Appl
29	46	59.0	675	3	US-08-947-965-76	Sequence 76, Appl
30	46	59.0	676	3	US-08-947-965-71	Sequence 71, Appl
31	46	59.0	680	3	US-08-947-965-77	Sequence 77, Appl
32	46	59.0	683	3	US-08-947-965-72	Sequence 2, Appl
33	46	59.0	687	3	US-08-947-965-75	Sequence 75, Appl
34	46	59.0	725	2	US-08-816-105A-1	Sequence 1, Appl
35	45	57.7	719	4	US-09-386-607-2	Sequence 11, Appl
36	44	56.4	14	1	US-08-204-656B-11	Sequence 11, Appl
37	44	56.4	14	1	US-08-470-702-11	Sequence 11, Appl
38	44	56.4	14	1	US-08-467-831-11	Sequence 11, Appl
39	44	56.4	454	1	US-07-930-686-10	Sequence 10, Appl
40	44	56.4	454	2	US-08-460-998-10	Sequence 10, Appl
41	44	56.4	833	4	US-09-514-302-3	Sequence 2, Appl
42	44	56.4	1938	4	US-09-514-302-2	Sequence 2, Appl
43	43	55.1	14	1	US-08-204-656B-12	Sequence 12, Appl
44	43	55.1	14	1	US-08-470-702-12	Sequence 12, Appl
45	43	55.1	14	1	US-08-467-831-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Bisgaard, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowrey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054,214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7
Query Match 100.0%; Score 78; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVHGYWOODIYSL 13
|||||
DB 78 AVHGYWOODIYSL 90

RESULT 2
US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-459-610-7

Query Match 100.0%; Score 78; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVHGYWOODIYSL 13
|||||
DB 78 AVHGYWOODIYSL 90

RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5830837disk of No. 5830837th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-343-804-7

Query Match 100.0%; Score 78; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVHGYWOODIYSL 13
|||||
DB 78 AVHGYWOODIYSL 90

RESULT 4
US-08-600-908A-10
; Sequence 10, Application US/08600908A
; Patent No. 5989169
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59891690 No. 5989169th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,908A
; FILING DATE: 13-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-600-908A-10

Query Match 100.0%; Score 78; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
|||||
DB 78 AYHGWOODIYSL 90

RESULT 5
US-08-683-838A-10
; Sequence 10, Application US/08683838A
; Patent No. 6022724
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 60227240 No. 6022724disk of No. 6022724th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,838A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/600,908
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394,204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-683-838A-10

Query Match 100.0%; Score 78; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AYHGWOODIYSL 13
|||||

Db 78 AYHGWOODIYSL 90

RESULT 6
US-09-182-859-7
; Sequence 7, Application US/09182859
; Patent No. 6143708
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgaard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796,204-US
; CURRENT APPLICATION NUMBER: US/09/182,859
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 0515/96
; EARLIER FILING DATE: 1996-04-30
; EARLIER APPLICATION NUMBER: 0712/96
; EARLIER FILING DATE: 1996-06-28
; EARLIER APPLICATION NUMBER: 0775/96
; EARLIER FILING DATE: 1996-07-11
; EARLIER APPLICATION NUMBER: 1263/96
; EARLIER FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 78; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
|||||
DB 78 AYHGWOODIYSL 90

RESULT 7
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michio
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POHMAN, HAIK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; STREET: G. Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Melsner, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 79.5%; Score 62; DB 2; Length 478;
Best Local Similarity 84.6%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGWOODIYSL 13
|||:|:|:|:|:|
Db 78 ATGTGWQDIYSL 90

RESULT 8
US-08-204-656B-2
Sequence 2, Application US/08204656B
Patent No. 5538882
GENERAL INFORMATION:
APPLICANT: Matsui, Ikuro
APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Miyairi, Sachio
APPLICANT: Honda, Koichi
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene of The Enzyme And Method For Producing
TITLE OF INVENTION: Oligosaccharide Using The Enzyme
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-204-656B-2

Query Match 69.2%; Score 54; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.34;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGWOODIYSL 13
|||:|:|:|:|:|
Db 79 AYHGFWMKNIYKI 91

RESULT 9
US-08-470-702-6
Sequence 6, Application US/08470702
Patent No. 5631149
GENERAL INFORMATION:
APPLICANT: MATSUI, IKURO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-6

Query Match 69.2%; Score 54; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.34;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGWOODIYSL 13
|||:|:|:|:|:|
Db 79 AYHGFWMKNIYKI 91

RESULT 10
US-08-467-831-6
; Sequence 6, Application US/08467831
; Patent No. 5635378
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,831
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-467-831-6

Query Match 69.2%; Score 54; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.34;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL.13
||||:|:|:|:
Db 79 AYHGFMKNITIKI 91

RESULT 11
US-08-204-656B-4
; Sequence 4, Application US/08204656B
; Patent No. 5518882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuo
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; APPLICANT: Honda, Koichi
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
; TITLE OF INVENTION: Oligosaccharide Using The Enzyme

; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,656B
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-204-656B-4

Query Match 67.9%; Score 53; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.49;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
||||:|:|:|:
Db 79 AYHGFMKNITIKI 91

RESULT 12
US-08-470-702-7
; Sequence 7, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-470-702-7

Query Match 67.9%; Score 53; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.49;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGWOODIYSL 13
||||:|:|:|:
Db 79 AYHGMMKNITYKI 91

RESULT 13
US-08-467-831-7
Sequence 7, Application US/08467831
Patent No. 5635378
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
NUMBER OF INVENTIONS: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,831
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-467-831-7

Query Match 67.9%; Score 53; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.49;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGWOODIYSL 13
||||:|:|:|:
Db 79 AYHGMMKNITYKI 91

RESULT 14
US-08-204-656B-6
Sequence 6, Application US/08204656B
Patent No. 5538882
GENERAL INFORMATION:
APPLICANT: Matsui, Ikuo
APPLICANT: Ishtkawa, Kazuhiko
APPLICANT: Miyairi, Sachio
APPLICANT: Honda, Koichi
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene Of The Enzyme And Method for Producing
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-204-656B-6

Query Match 64.1%; Score 50; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 1.5;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AYHGWOODIYSL 13
||||:|:|:|:
Db 79 AYHGMMKNITYKI 91

RESULT 15
US-08-470-702-8
; Sequence 8, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 810 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-470-702-8

Query Match 64.1%; Score 50; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 1.5;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 AYHGTWQDDIYSL 13
|||||::|||:
Db 79 AYHGLMKNIYKI 91

Search completed: June 13, 2002, 08:54:29
Job time: 314 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:06 ; Search time 125.81 Seconds
(without alignments)
17.876 Million cell updates/sec

Title: US-09-710-339-2_COPY_98_110
Sequence: 1 AHHGYWQDIYSL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTRMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	499	3	Q96TH4 aspergillus
2	75	96.2	490	3	Q9UV07 emericella
3	68	87.2	128	5	015751 dictyostell
4	66	84.6	631	3	Q92394 cryptococcu
5	64	82.1	507	3	Q08806 debaryomyce
6	63	80.8	640	3	013296 aspergillus
7	62	79.5	624	3	001117 lipomyces k
8	59	75.6	482	2	060051 thermocactin
9	54	69.2	491	3	013996 schizosacch
10	54	69.2	623	3	Q9UV09 emericella
11	53	67.9	564	3	Q9Y7S9 schizosacch
12	52	66.7	711	16	Q99ZB3 streptococc
13	50	64.1	625	3	074922 schizosacch
14	49	62.8	690	16	Q9KLB6 vibrio chol
15	49	62.8	713	2	Q9FSW3 bacillus ci
16	47	60.3	483	16	Q9RUB8 deinococcus

17	46	59.0	692	2	Q30565 bacillus br
18	46	59.0	704	2	Q82984 bacillus sp
19	46	59.0	711	2	Q9ZAQ0 bacillus st
20	46	59.0	714	2	Q52766 paenibacill
21	46	59.0	724	2	Q9K515 actinoplan
22	46	59.0	725	2	Q59239 bacillus sp
23	45	57.7	739	1	Q9UWN2 thermococcu
24	44	56.4	459	5	Q45132 haemophilus
25	44	56.4	677	2	Q9RHR1 klebsiella
26	44	56.4	1938	2	P70983 bacillus sp
27	43	55.1	134	5	Q9MSU9 drosophila
28	43	55.1	158	5	Q965N4 caenorhabdi
29	43	55.1	464	4	Q9NSY6 mus sapien
30	43	55.1	476	11	Q99LV4 mus musculu
31	43	55.1	552	10	Q9LVM8 arabidopsis
32	43	55.1	612	4	Q96K36 homo sapien
33	43	55.1	768	4	Q96K54 homo sapien
34	43	55.1	793	4	Q96K44 homo sapien
35	43	55.1	793	4	Q96K38 homo sapien
36	43	55.1	793	11	Q9DC23 mus musculu
37	42	53.8	198	5	Q9VRM1 drosophila
38	42	53.8	275	2	Q9ALM7 saccharopol
39	42	53.8	397	16	Q913T2 pseudomonas
40	41	52.6	153	2	Q93A49 enterococcu
41	41	52.6	410	13	Q9PW33 typhloocte
42	41	52.6	413	10	Q9M6R9 malus domes
43	41	52.6	517	2	Q9X8B6 streptomyce
44	41	52.6	1021	2	Q9K5L6 actinoplan
45	41	52.6	1798	2	Q9K211 streptomyce

ALIGNMENTS

RESULT 1
Q96TH4 ID Q96TH4 PRELIMINARY; PRT; 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
GN AMYA OR AMY1.
OS Aspergillus oryzae, and
OC Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062, 5059;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A.oryzae; STRAIN=R1840;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator
gene, amyR, involved in the amylolytic gene expression in Aspergillus
oryzae."
RL Biosci. Biotechnol. Biochem. 64:816-827(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=A.flavus; STRAIN=86-10D;
RA Fakhoury A.M., Woloshuk C.P.;
RT "Amyl, the alpha-amylase gene of Aspergillus flavus: Involvement in
afatoxin biosynthesis in maize kernels."
RL Phytopathology 89:908-914(1999).
DR EMBL; AB021876; BA85703.1; -;
DR EMBL; AF139925; AAF14264.1; -;
SQ SEQUENCE 499 AA; 54804 MW; EEFA2ADA71D20DA9 CRC64;
Query Match 100.0%; Score 78; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

	Query Match	87.2%	Score 68:	DB 5;	Length 128;
	Best Local Similarity	83.3%;	Pred. No.	0.0014;	
	Matches 10; Conservative	2;	Mismatches	0;	Gaps 0;
QY	2 YHGWOODIRSL 13				
	:				
Dd	3 YHGYWOODITV 14				

ID	008806	PRELIMINARY:	PRT:	507 AA.
AC	008806;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	ALPHA-AMYLASE 2 (EC 3.2.1.1) (1.4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).			
GN	SWA2.			
OS	Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Debaryomyces.			
OX	NCBI_TaxID=27300;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 26077.			
RX	MEDLINE=93365041; PubMed=8359835.			
RA	Clares M.G., Abarta D., Fernandez-Iobato M., Jimenez A.;			
RT	Molecular structure of the SWA2 gene encoding an Amy1-related alpha			
RT	amylase from Schwanniomyces occidentalis.*;			
RL	Curr. Genet. 24:75-83(1993).			
CC	-1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC			
CC	LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.			
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY			
CC	-1- ENZYME REGULATION: INDUCES EXPRESSION UNDERLIES CATABOLITE			
CC	REPRESSION BY GLUCOSE.			

CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL: X73497; CA51912.1; -.
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha_amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KM Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KM Glycoprotein.
 FT ACT_SITE 238 238 BY SIMILARITY.
 FT ACT_SITE 242 242 BY SIMILARITY.
 FT ACT_SITE 329 329 BY SIMILARITY.
 FT DISULFID 62 70 BY SIMILARITY.
 FT DISULFID 182 196 BY SIMILARITY.
 FT DISULFID 272 315 BY SIMILARITY.
 FT DISULFID 470 505 BY SIMILARITY.
 FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (PROBABLE).
 SQ SEQUENCE 507 AA; 55966 MW; 3A562E95BD8AD63 CRC64;

Query Match 82.1%; Score 64; DB 3; Length 507;
 Best Local Similarity 69.2%; Pred. No. 0.026;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGWMODIYSL 13
 DB 110 AYHGWMODIYSL 122

RESULT 6
 O13296 PRELIMINARY; PRT; 640 AA.
 ID O13296;
 AC O13296;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ACID-STABLE ALPHA-AMYLASE.
 OS Aspergillus kawachii.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OX NCBI_TaxID=40384;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kaneko A.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kaneko A.; Sudo S.; Sakamoto Y.; Tamura G.; Ishikawa T.; Ohba T.;
 RT "Molecular-cloning and determination of the nucleotide-sequence of a
 RT gene encoding an acid-stable alpha-amylase from Aspergillus-
 RT kawachii.";
 RL U. Ferment. Bioeng. 81:292-298(1996).
 DR EMBL: AB008370; BAA22993.1; -.
 DR HSSP: P56271; 2AAA.
 DR InterPro: IPR000461; Alpha_amylase.
 DR InterPro: IPR002044; CBD_4.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR ProDom: PD001568; CBD_4; 1.
 SQ SEQUENCE 640 AA; 69507 MW; 676BF9D0236720DD CRC64;

Query Match 80.8%; Score 63; DB 3; Length 640;
 Best Local Similarity 90.9%; Pred. No. 0.048;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGWMODIYSL 11
 DB 99 AYHGWMODIYSL 109

RESULT 7
 Q01117 PRELIMINARY; PRT; 624 AA.
 ID Q01117

AC Q01117;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1)
 DE (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE) (UKAI).
 GN UKAI.
 OS Lipomyces kononenkoae.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Lipomycetaceae; Lipomyces.
 OX NCBI_TaxID=34357;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IGC4052B;
 RA MEDLINE=96105202; PubMed=8529895;
 RX Steyn A.J.C.; Marmur J.; Pretorius I.S.;
 RT "Cloning, sequence analysis and expression in yeasts of a cDNA
 RT containing a Lipomyces kononenkoae alpha-amylase-encoding gene.";
 RL Gene 166:65-71(1995).
 RN [2]
 RP SEQUENCE OF 29-44.
 RC STRAIN=IGC4052B;
 RX MEDLINE=96132108; PubMed=8593683;
 RA Steyn A.J.C.; Pretorius I.S.;
 RT "Characterization of a novel alpha-amylase from Lipomyces kononenkoae
 RT and expression of its gene (UKAI) in Saccharomyces cerevisiae.";
 RL Curr. Gene. 28:526-533(1995).
 CC -1- FUNCTION: LIBERATES REDUCING GROUPS FROM POLYMERS CONTAINING BOTH
 CC 1,4-ALPHA AND 1,6-ALPHA BONDS.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL: U30376; AAC49622.1; ALT_INIT.
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha_amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KM Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Calcium;
 KM Glycoprotein.
 FT SIGNAL 1 28
 FT CHAIN 29 624 ALPHA-AMYLASE.
 FT ACT_SITE 353 353 BY SIMILARITY.
 FT ACT_SITE 357 357 BY SIMILARITY.
 FT ACT_SITE 377 377 BY SIMILARITY.
 FT ACT_SITE 444 444 BY SIMILARITY.
 FT DISULFID 177 185 BY SIMILARITY.
 FT DISULFID 297 311 BY SIMILARITY.
 FT DISULFID 387 430 BY SIMILARITY.
 FT DISULFID 587 622 BY SIMILARITY.
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 624 AA; 68876 MW; 87EB16534F5A99F CRC64;

Query Match 79.5%; Score 62; DB 3; Length 624;
 Best Local Similarity 69.2%; Pred. No. 0.067;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGWMODIYSL 13
 DB 225 AYHGWMODIYSL 237

RESULT 8
 Q60051 PRELIMINARY; PRT; 482 AA.
 ID Q60051;
 AC Q60051;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1).
 GN AMYV.
 OS Thermactinomyces vulgaris.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Thermactinomyces.
 OX NCBI_TaxId=2026;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-K94;
 RX MEDLINE=95031040; PubMed=7944369;
 RA Hofmeister B., Koenig S., Hoang V., Engel J., Mayer G., Hansen G.,
 RT "The gene amy(TY1) codes for a nonlucogenic alpha-amylase from
 Thermactinomyces vulgaris 94-2A in Bacillus subtilis."
 RL Appl. Environ. Microbiol. 60:381-389(1994).
 DR EMBL: X69807; CAA49465.1; -.
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase-C; 1.
 KW Signal; Hydrolase; Glycosidase.
 FT SIGNAL
 FT CHAIN 30 482
 SQ SEQUENCE 482 AA; 54483 MW; D8F98C370B579025 CRC64;

Query Match 75.6%; Score 59; DB 2; Length 482;
 Best Local Similarity 69.2%; Pred. No. 0.16;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYHGYYQODIYSL 13
 Db 107 AYHGYYTDFYGSV 119
 ||||| |||

RESULT 9
 ID 013996 PRELIMINARY; PRT; 491 AA.
 AC 013996;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PUTATIVE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1).
 GN SPAC27E2.O1.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.,
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL: Z69878; CAB11675.1; -.
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
 KW Glycoprotein; Signal.
 FT SIGNAL
 FT CHAIN 1 491
 FT ACT_SITE 221 491
 FT ACT_SITE 225 491
 FT ACT_SITE 313 491
 FT ACT_SITE 313 491
 FT CARBOHYD 307 491
 FT CARBOHYD 319 491
 FT CARBOHYD 391 491
 SQ SEQUENCE 491 AA; 56678 MW; 2013808DC0EE581D CRC64;

Query Match 69.2%; Score 54; DB 3; Length 491;
 Best Local Similarity 61.5%; Pred. No. 1;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYHGYYQODIYSL 13
 Db 94 AYHGYYWEDLTOL 106
 ||||| :|

RESULT 10
 ID 09V09 PRELIMINARY; PRT; 623 AA.
 AC 09V09;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ALPHA-AMYLASE.
 GN AMYB.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; Emericella.
 OX NCBI_TaxId=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boase N.A., Murphy R.L., Kelly J.M.;
 RT "Characterisation and regulation of an alpha-amylase gene in
 Aspergillus nidulans."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF208224; AAF17100.1; -.
 DR HSSP: P56271; 2AAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR PRINTS: PR00110; ALPHAAMYLASE.
 DR PRODOM: PD001566; CBD_4; 1.
 SQ SEQUENCE 623 AA; 68598 MW; 822F002C37F5A9A4 CRC64;

Query Match 69.2%; Score 54; DB 3; Length 623;
 Best Local Similarity 58.3%; Pred. No. 1.3;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 YHGYYQODIYSL 13
 Db 98 YHGYYKNYGSV 109
 :|||||:|

RESULT 11
 ID 09Y7S9 PRELIMINARY; PRT; 564 AA.
 AC 09Y7S9;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PROBABLE ALPHA-AMYLASE C63.02C PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-
 DE GLUCAN GLUCANOHYDROLASE).
 GN SPOC63.02C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.,
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
 CC SIMILARITY).


```

CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR EMBL: AL049522; CAB40006.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KW Calcium; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 564
FT DOMAIN 537 540
FT ACT_SITE 229 229
FT ACT_SITE 322 322
FT DISULFID 51 59
FT DISULFID 172 188
FT DISULFID 263 306
FT CARBOHYD 181 181
FT CARBOHYD 235 235
FT CARBOHYD 282 282
FT CARBOHYD 305 305
FT CARBOHYD 438 438
FT CARBOHYD 447 447
FT CARBOHYD 498 498
SQ SEQUENCE 564 AA; 63205 MW; 1229FD4AEC702FED0 CRC64;

Query Match
Best Local Similarity 67.9%; Score 53; DB 3; Length 564;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AYHGWMODIYSL 13
    ||||| 1 11
DB 100 AYHGWMNTDESL 112

RESULT 12
Q99ZB3 PRELIMINARY; PRT; 711 AA.
AC Q99ZB3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE CYCLOMALTODEXTRIN GLUCANOTRANSFERASE (EC 2.4.1.19).
GN AMYA OR SPY1302.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_Taxid=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL: AE006569; AAK3149.1; -.
DR HSSP: P4379; 1CDG.
DR InterPro: IPR00461; Alpha-amylase.
DR InterPro: IPR002453; Beta-tubulin.
DR InterPro: IPR002044; CBD_4.
DR InterPro: IPR002909; IPT_TIG.
DR Pfam: PF00128; alpha-amylase; 1.
DR Pfam: PF02806; alpha-amylase; 1.
DR Pfam: PF00686; CBD_4; 1.
DR Pfam: PF01833; TIG; 1.
DR PRINTS: PR00110; ALPHAAMYLASE.
DR PRODOM: PD001568; CBD_4; 1.
DR PROSITE: PS00228; TUBULIN_B-ATOREG; UNKNOWN_1.
DR Transferrase; Glycosyltransferase; Complete proteome.
SQ SEQUENCE 711 AA; 79544 MW; 88F7AA9424C0E717 CRC64;

```

```

Query Match
Best Local Similarity 66.7%; Score 52; DB 16; Length 711;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGWMODIY 11
    ||||| 1 11
DB 122 AYHGWMNDFF 132

RESULT 13
Q74922 PRELIMINARY; PRT; 625 AA.
AC Q74922;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE.
GN SPC0757.12.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Lyne M., Rajadream M.A., Barrell B.G., Bothe G., Pohl T.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL031825; CAA21237.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
SQ SEQUENCE 625 AA; 68087 MW; 30A58CE6D6BC07B3 CRC64;

Query Match
Best Local Similarity 64.1%; Score 50; DB 3; Length 625;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AYHGWMODIYSL 13
    ||||| 1 11
DB 100 AYHGWMNTDFQL 112

RESULT 14
Q9KL86 PRELIMINARY; PRT; 690 AA.
AC Q9KL86;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE.
GN VCA0860.
OS Vibrrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_Taxid=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL: AE004413; AAF96758.1; -.

```

DR TIGR: VCA0860; -
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Complete proteome.
 SQ SEQUENCE 690 AA; 77331 MW; 7C82C00C928FDE41 CRC64;

Query Match 62.8%; Score 49; DB 16; Length 690;
 Best Local Similarity 77.8%; Pred. No. 9.1;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGYWOOD 9
 ||||| :!
 Db 288 AYHGYWRD 296

RESULT 15

O9F5W3 PRELIMINARY; PRT; 713 AA.
 AC O9F5W3.
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE CYCLODEXTRIN GLUCANOSYLTRANSFERASE.
 OS Bacillus circulans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1397;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-All;
 RA Rhiphanitchayakit V., Tonozuka T., Sakano Y.;
 RT "Cloning of Cyclodextrin Glucanotransferase Gene from Bacillus
 RT circulans All.";
 RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF302787; AAC31622.1; -.
 DR HSSP: P05618; 1PAM.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR PRINTS: PR00110; ALPHAAMYLASE.
 DR ProDom: PD001568; CBD_4; 1.
 KW Transferrase.
 SQ SEQUENCE 713 AA; 78271 MW; 4FF365E1F8F93BAD CRC64;

Query Match 62.8%; Score 49; DB 2; Length 713;
 Best Local Similarity 77.8%; Pred. No. 9.4;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGYWOOD 9
 ||||| :!
 Db 123 AYHGYWRD 131

Search completed: June 13, 2002, 09:11:07
 Job time: 1052 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:48 ; Search time 38.1 Seconds

(without alignments)
13.211 Million cell updates/sec

Title: us-09-710-339-2_COPY_98_110
Perfect score: 78
Sequence: 1 AYHGYWQDIYSL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	78	100.0	498	1 AMYA_ASPAW	Q02905 aspergillus
2	78	100.0	499	1 AMYA_ASPOR	P10529 aspergillus
3	78	100.0	499	1 AMYA_ASPAW	Q02906 aspergillus
4	78	100.0	499	1 AMY_ASPSH	P30282 aspergillus
5	67	85.9	513	1 AMY2_SCHPO	O14154 schizosacch
6	64	82.1	484	1 AMYA_ASPNG	P56271 aspergillus
7	58	74.4	494	1 AMY1_SACFI	P21567 saccharomyc
8	56	71.8	478	1 YDOL_SCHPO	O10427 schizosacch
9	52	66.7	581	1 AMY1_SCHPO	Q09840 schizosacch
10	50	64.1	512	1 AMY1_DEROC	P19269 debaromyces
11	49	62.8	528	1 AMY_BACCT	P08137 bacillus ci
12	49	62.8	712	1 CDGT_BACSS3	P09121 bacillus sp
13	49	62.8	713	1 AMYR_BACSS8	P17692 bacillus sp
14	49	62.8	713	1 CDG2_PAPMA	P31835 paenibacill
15	49	62.8	713	1 CDGT_BACSO	P05618 bacillus sp
16	49	62.8	713	1 CDGT_BACSP	P30921 bacillus sp
17	49	62.8	713	1 CDGT_BACCI	P43379 bacillus ci
18	49	62.8	718	1 CDGT_BACCI	P30920 bacillus ci
19	49	62.8	718	1 CDGT_BACLI	P14014 bacillus li
20	49	62.8	718	1 CDGT_BACSS	P31747 bacillus sp
21	47	60.3	655	1 CDGT_KLEPN	P08704 klebsiella
22	46	59.0	703	1 CDGT_BACSS2	P31746 bacillus sp
23	46	59.0	704	1 CDGT_BACOH	P27036 bacillus oh
24	46	59.0	710	1 CDGT_THETU	P26827 thermomater
25	46	59.0	711	1 CDGT_BACST	P31797 bacillus st
26	46	59.0	714	1 CDG1_PAPMA	P04830 paenibacill
27	45	57.7	717	1 AMYM_BACST	P19531 bacillus st
28	44	56.4	676	1 AMY1_ECOLI	P25728 escherichia
29	44	56.4	2158	1 MY9B_HUMAN	Q13459 homo sapien
30	43	55.1	1196	1 AMYB_PABPO	P21543 paenibacill
31	41	52.6	499	1 MYFD_RHOCA	P08717 rhodobacter
32	41	52.6	919	1 AMY_STRLI	O05884 streptomyces
33	40.5	51.9	604	1 RCOL_NEUCR	P78706 neurospora

34	40	51.3	508	1	GSN1 MEDTR	Q92NX6 medicago tr
35	40	51.3	734	1	PSAB_CYACA	Q9L1Q6 cyanidium c
36	40	51.3	1210	1	YEAH_ECOLI	P33546 escherichia
37	39	50.0	395	1	RRP_PNDVA	P16073 newcastle d
38	39	50.0	395	1	RRP_PNDVB	P24698 newcastle d
39	39	50.0	395	1	RRP_PNDV2	P06427 newcastle d
40	38	48.7	383	1	PEPE_CHICK	P16476 gallus gall
41	38	48.7	387	1	PEP3_RABIT	P27822 oryctolagus
42	38	48.7	413	1	HMT1_DROME	P02835 drosophila
43	38	48.7	501	1	C6AH_DROME	Q9V770 drosophila
44	38	48.7	502	1	C6AN_DROME	Q9V771 drosophila
45	38	48.7	559	1	CYST_THIRO	P52673 thiocapsa r

ALIGNMENTS

RESULT	ID	AMYA_ASPAW	STANDARD	PRT	498 AA.
AC	002905				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-OCT-1996	(Rel. 34, Last annotation update)			
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan				
DE	glucanohydrolase A).				
GN	AMYA.				
OS	Aspergillus awamori.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_TaxID=105351;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=UW143P.				
RX	MEDLINE=90254827; PubMed=2340591;				
RA	Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,				
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,				
RA	Berka R.M.;				
RT	"Cloning, characterization, and expression of two alpha-amylase genes				
RT	from Aspergillus niger var				
RT	curr. Genet. 17:203-212(1990)."				
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic				
CC	linkages in oligosaccharides and polysaccharides.				
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO				
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; X52755; CAA36966.1; -				
DR	HSSP; P10529; TTA.				
DR	InterPro: IPR000461; Alpha.amylase.				
DR	Pfam: PF00128; alpha-amylase; 1.				
KW	Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;				
KW	glycoprotein; Signal; Multigene family.				
FT	SIGNAL	1	21	POTENTIAL.	
FT	CHAIN	22	498	ALPHA-AMYLASE A.	
FT	ACT_SITE	227	227	BY SIMILARITY.	
FT	ACT_SITE	231	231	BY SIMILARITY.	
FT	ACT_SITE	251	251	BY SIMILARITY.	
FT	ACT_SITE	318	318	BY SIMILARITY.	
FT	DISULFID	51	59	BY SIMILARITY.	
FT	DISULFID	171	185	BY SIMILARITY.	
FT	DISULFID	261	304	BY SIMILARITY.	
FT	DISULFID	461	496	BY SIMILARITY.	
FT	CARBOHYD	218	218	N-LINKED (GLCNAC...) (POTENTIAL).	
SQ	SEQUENCE	498 AA:	54880 MM:	7658511BC01ABA01 CRC64;	

Query Match 100.0%; Score 78; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. NO. 1.4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
 |||||||
 DB 99 AYHGWOODIYSL 111

RESULT 2
 AMYA_ASPOR STANDARD; PRT: 499 AA.

ID AMYA_ASPOR
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 DE (1,4-alpha-D-glucan glucohydrolase).
 GN AMY1 AND AMY2 AND AMY3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
 OC Eurotiaceae; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 63303;
 RA MEDLINE=89237897; PubMed=2785629;
 RA Wirsel S., Lachmund A., Wilhardt G., Rutkowski E.;
 RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT Intron-exon organization.";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89378767; PubMed=2789162;
 RA Genes M.J., Dove M.J., Seliy V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns.";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshikawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae.";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90128276; PubMed=2612911;
 RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
 RA Ueda S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes.";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A.";
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RA MEDLINE=74001521; PubMed=4733850;
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RT A with trypsin and chymotrypsin.";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; PubMed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution.";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84212370; PubMed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A.";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; PubMed=9283074;
 RA Brzozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution.";
 RL Biochemistry 36:10837-10845(1997).
 CC -1 CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -1 COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1 MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04348; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JK0240; JK0240.
 DR PIR: JTM0466; JTM0466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR GlycosultedB: P10529;
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase: 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT CONFLICT 93 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTID=CAR_000125.
 Q -> R (IN AMY3).
 F -> L (IN AMY3).
 TP -> DC (IN REF. 5).
 Q -> T (IN REF. 5).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 G -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).

FT CONFLICT 370 370 L -> A (IN REF. 4).
 FT CONFLICT 406 409 WPLY -> PYI (IN REF. 5).
 FT CONFLICT 448 448 G -> S (IN REF. 5).
 FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).
 FT HELIX 24 27
 FT TURN 28 29
 FT STRAND 32 35
 FT HELIX 37 40
 FT STRAND 42 42
 FT TURN 43 44
 FT HELIX 44 44
 FT STRAND 53 55
 FT HELIX 61 61
 FT HELIX 63 67
 FT TURN 68 69
 FT HELIX 70 74
 FT TURN 75 77
 FT STRAND 80 83
 FT STRAND 87 89
 FT STRAND 94 94
 FT TURN 95 96
 FT STRAND 97 97
 FT TURN 100 101
 FT STRAND 105 111
 FT TURN 113 115
 FT TURN 118 119
 FT HELIX 120 130
 FT TURN 131 132
 FT STRAND 134 139
 FT STRAND 143 143
 FT STRAND 146 146
 FT HELIX 150 152
 FT TURN 155 156
 FT STRAND 159 159
 FT HELIX 164 166
 FT STRAND 167 167
 FT STRAND 172 172
 FT HELIX 179 184
 FT STRAND 186 188
 FT STRAND 193 194
 FT STRAND 196 197
 FT TURN 199 200
 FT HELIX 202 219
 FT TURN 220 220
 FT STRAND 223 226
 FT TURN 227 228
 FT HELIX 229 231
 FT TURN 234 235
 FT HELIX 236 244
 FT STRAND 247 250
 FT HELIX 257 265
 FT TURN 266 266
 FT STRAND 269 271
 FT HELIX 273 283
 FT TURN 286 287
 FT HELIX 290 303
 FT STRAND 307 309
 FT TURN 311 312
 FT TURN 317 318
 FT HELIX 322 325
 FT HELIX 329 341
 FT STRAND 345 349
 FT TURN 350 351
 FT HELIX 352 354
 FT TURN 355 355
 FT TURN 361 364
 FT HELIX 368 371
 FT TURN 372 372

Query Match 100.0%; Score 78; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGYWOODIYSL 13

DB 99 AYHGYWOODIYSL 111

RESULT 3
 AMYB_ASPAW STANDARD; PRT; 499 AA.

AC 002906;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Alpha-amyase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 DE glucanohydrolase B).
 GN AMYB.
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=105351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UVK143F;
 RX MEDLINE=90254827; PubMed=2340591;
 RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
 RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
 RA Berka R.M.;

RT "Cloning, characterization, and expression of two alpha-amyase genes
 from Aspergillus niger var. awamori.";
 RL Curr. Genet. 17:203-212(1990).
 CC -I CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X52756; CA36967.1; -.

DR HSPD; P10529; 7TAA.

DR InterPro; IPR000461; Alpha_amyase.

DR Pfam; PF00128; alpha-amyase; 1.

KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 499 ALPHA-AMYLASE B.

FT ACT_SITE 227 227 BY SIMILARITY.

FT ACT_SITE 231 231 BY SIMILARITY.

FT ACT_SITE 251 251 BY SIMILARITY.

FT ACT_SITE 318 318 BY SIMILARITY.

FT DISULFID 51 59 BY SIMILARITY.

FT DISULFID 171 185 BY SIMILARITY.

FT DISULFID 261 304 BY SIMILARITY.

FT DISULFID 461 496 BY SIMILARITY.

FT CARBOHYD 218 218 N-LINKED (GLCNAC...) (POTENTIAL).

SO SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;

Query Match 100.0%; Score 78; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGYWOODIYSL 13

DB 99 AYHGYWOODIYSL 111

RESULT 4

AMY_ASPSH

```

ID AMY_ASPESH STANDARD: PRT: 499 AA.
AC P30292;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylose precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
  glucanohydrolase).
GN AMY.
OS Aspergillus shirousami.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5070;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92323146; PubMed=1368777;
RA Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT "Cloning of the alpha-amylose cDNA of Aspergillus shirousami and its
  expression in Saccharomyces cerevisiae."
RL Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC -1 CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
  linkages in oligosaccharides and polysaccharides.
CC -1 COPFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D10461; BAA01255.1; -.
CC PIR: J50663; J50663.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylose.
DR Pfam: PF00128; alpha-amylose; 1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KM Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 1 21
FT ACT_SITE 22 499 ALPHA-AMYLASE.
FT ACT_SITE 227 227
FT ACT_SITE 231 231 BY SIMILARITY.
FT ACT_SITE 251 251 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 51 59
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 261 304 BY SIMILARITY.
FT DISULFID 461 496 BY SIMILARITY.
FT CARBOHYD 218 218 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ SEQUENCE 499 AA: 54852 MW: 1567450DA01C03F CRC64;

Query Match 100.0%; Score 78; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 1,4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AYHGWOODIYSL 13
DB 99 AYHGWOODIYSL 111

```

```

OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RA Skellton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
  linkages in oligosaccharides and polysaccharides.
CC -1 COPFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
  SIMILARITY).
CC -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z98762; CAB1471.1; -.
CC HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylose.
DR Pfam: PF00128; alpha-amylose; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KM Calcium; Glycoprotein; Signal.
FT SIGNAL 1 25
FT CHAIN 1 25
FT ACT_SITE 26 513 PROBABLE ALPHA-AMYLASE C4A8.01.
FT ACT_SITE 226 226 BY SIMILARITY.
FT ACT_SITE 230 230 BY SIMILARITY.
FT ACT_SITE 250 250 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 52 60
FT DISULFID 171 184 BY SIMILARITY.
FT DISULFID 260 304 BY SIMILARITY.
FT DISULFID 454 488 BY SIMILARITY.
FT DISULFID 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 513 AA: 58715 MW: 455DD9FEA428C182 CRC64;

Query Match 85.9%; Score 67; DB 1; Length 513;
Best Local Similarity 76.9%; Pred. No. 0.00096;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 AYHGWOODIYSL 13
DB 100 AYHGWOODIYSL 112

```

RESULT 5
 AMY2_SCHPO
 ID AMY2_SCHPO STANDARD: PRT: 513 AA.
 AC 014154;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Probable alpha-amylose C4A8.01 precursor (EC 3.2.1.1) (1,4-alpha-D-
 glucan glucanohydrolase).
 GN SPAC4A8.01.

RESULT 6
 AMYA_ASPENG
 ID AMYA_ASPENG STANDARD: PRT: 484 AA.
 AC P56271;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Acid alpha-amylose (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RX MEDLINE=91002514; PubMed=2207069;
 RA Boel E., Brady L., Brzozowski A.M., Derewenda Z., Dodson G.G.,
 Jensen V.J., Petersen S.B., Swift H., Thim L., Woldike H.F.;
 FT "Calcium binding in alpha-amyloses: an X-ray diffraction study at

RT 2.1-A resolution of two enzymes from *Aspergillus*."
 RL Biochemistry 29:6244-6249(1990).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR PDB: 2AAV; 15-JUL-93.
 DR InterPro: IPR000461; Alpha-amylase.
 KW Pfam: PF00128; alpha-amylase; 1.
 KM Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
 KM Glycoprotein; 3D-structure.
 FT ACT_SITE 206 206
 FT ACT_SITE 210 210
 FT ACT_SITE 230 230
 FT ACT_SITE 297 297
 FT DISULFID 30 38
 FT DISULFID 150 164
 FT DISULFID 240 283
 FT DISULFID 440 475
 FT CARBOHYD 24 24
 FT CARBOHYD 157 157
 FT CARBOHYD 197 197
 FT CARBOHYD 197 197
 SQ SEQUENCE 484 AA; 52935 MW; 04D596E34680656D CRC64;

Query Match 82.1%; Score 64; DB 1; Length 484;
 Best Local Similarity 76.9%; Pred. No. 0.0028;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVHGYWQDIIYSL 13
 ||||| :
 DB 78 AVHGYWQKIDYV 90

RESULT 7
 AM1_SACFI STANDARD: PRT; 494 AA.
 ID P21567;
 AC 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Alpha-amylase precursor (EC 3.2.1.1).
 GN ALP1.
 OS Saccharomycopsis fibuligera (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
 OX NCBI_TaxID=4944;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87276512; PubMed=3497057;
 RA Itoh T., Yamashita I., Fukui S.;
 RT "Nucleotide sequence of the alpha-amylase gene (ALP1) in the yeast
 RL *Saccharomycopsis fibuligera*."
 RL FEBS Lett. 219:339-342(1987).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: X05791; CAA29233.1; -.
 DR PIR: S00064; ALBYAF.
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.

DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase.C; 1.
 KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
 KM Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 494
 FT ACT_SITE 233 233
 FT ACT_SITE 237 237
 FT ACT_SITE 324 324
 FT CARBOHYD 224 224
 FT DISULFID 57 65
 FT DISULFID 177 191
 FT DISULFID 267 310
 FT DISULFID 462 493
 SQ SEQUENCE 494 AA; 54387 MW; 7E7DBFDBE58B67 CRC64;

Query Match 74.4%; Score 58; DB 1; Length 494;
 Best Local Similarity 61.5%; Pred. No. 0.028;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVHGYWQDIIYSL 13
 ||||| :
 DB 105 AVHGYWQKNIYKI 117

RESULT 8
 YD1_SCHPO STANDARD: PRT; 478 AA.
 ID Q10427; Q9UNO;
 AC 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 55.4 kDa protein C188.01c in chromosome III.
 GN SPEC188.01c OR SPEC11E10.09c.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=972.
 RA Ramsperger U., Pohl T., Wood V., Rajandream M.A., Barrell B.G.;
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AL121783; CAB57851.1; -.
 DR EMBL: AL049662; CAB41221.1; -.
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hypothetical protein; Hydrolase; Glycosidase.
 FT ACT_SITE 216 216
 FT ACT_SITE 220 220
 FT ACT_SITE 308 308
 FT ACT_SITE 308 308
 SQ SEQUENCE 478 AA; 55442 MW; 3C65EAD005DF7D3E CRC64;

Query Match 71.8%; Score 56; DB 1; Length 478;
 Best Local Similarity 69.2%; Pred. No. 0.058;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AVHGYWQDIIYSL 13

Db 89 AYHGWMQDMTOL 101

```

RESULT 9
AMYL_SCHPO STANDARD: PRT: 581 AA.
ID AMYL_SCHPO
AC Q09840:
DE 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Probable alpha-amylase C23D3.14C precursor (EC 3.2.1.1) (1,4-alpha-D-
  glucan glucohydrolase).
GN SPAC23D3.14C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
  Schizosaccharomycetales; Schizosaccharomycetaceae;
  Schizosaccharomyces.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Niblett D., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
  Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RL CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
  linkages in oligosaccharides and polysaccharides
  -I- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
  SIMILARITY).
  -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL: Z64354; CAA91249.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
  Calcium; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 581
FT ACT_SITE 233 233 BY SIMILARITY.
FT ACT_SITE 237 237 BY SIMILARITY.
FT ACT_SITE 325 325 BY SIMILARITY.
FT DISULFID 176 64 BY SIMILARITY.
FT DISULFID 267 311 BY SIMILARITY.
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 581 AA; 67004 MW; FE9DE9D323E1890 CRC64;

```

Query Match 66.78; Score 52; DB 1; Length 581;
 Best Local Similarity 53.84; Pred. No. 0.33;
 Matches 7; Conservativity 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYHGWMQDIYSL 13
 DB 104 SYHGWTNTELEL 116

RESULT 10
 AMYL_DEBOC STANDARD: PRT: 512 AA.
 AC P19269;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase 1 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
  glucohydrolase).
GN AMY1.
OS Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_Taxid=27300;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 26077;
RX MEDLINE=90032659; PubMed=2806251;
RA Strasser A.W.M., Selk R., Dohmen R.J., Niermann T., Bielefeld M.,
  Seeboth P., Tu G., Hollenberg C.P.;
  "Analysis of the alpha-amylase gene of Schwanniomyces occidentalis
  and the secretion of its gene product in transformants of different
  yeast genera."
  Eur. J. Biochem. 184:699-706(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CCRC 21164;
RX MEDLINE=92120467; PubMed=1769525;
RA Wu F.M., Wang T.T., Hsu W.H.;
  "The nucleotide sequence of Schwanniomyces occidentalis alpha-amylase
  gene."
  FEBS Microbiol. Lett. 66:313-318(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 26077 / CBS 2863;
RX MEDLINE=92307400; PubMed=612414;
RA Park J.C., Bai S., Tai C.Y., Chun S.B.;
  "Nucleotide sequence of the extracellular alpha-amylase gene in the
  yeast Schwanniomyces occidentalis ATCC 26077."
  FEBS Microbiol. Lett. 72:17-23(1992).
RL CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
  linkages in oligosaccharides and polysaccharides.
  -I- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
  -I- ENZYME REGULATION: ALPHA-AMYLASE EXPRESSION UNDERLIES CATABOLITE
  REPRESSION BY GLUCOSE.
CC
CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL: S77586; AAB21151.2; -.
DR EMBL: X16040; CAA34162.1; -.
DR EMBL: X62079; CAA43995.1; -.
DR EMBL: S38381; AAB22383.2; -.
DR PIR: S06115; S06115.
DR PIR: S23355; S23355.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
  Glycoprotein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 512
FT ACT_SITE 242 242 BY SIMILARITY.
FT ACT_SITE 246 246 BY SIMILARITY.
FT ACT_SITE 333 333 BY SIMILARITY.
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (PROBABLE).
FT DISULFID 66 74 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 276 319 BY SIMILARITY.
FT DISULFID 475 510 BY SIMILARITY.
FT VARIANT 32 32 M->K (IN STRAINS CCRC 21164 AND ATCC
  26077).

```



```

FT VARIANT 36 36 S -> G (IN STRAIN CCRC 21164).
FT VARIANT 73 73 Y -> I (IN STRAIN ATCC 26077).
FT VARIANT 280 280 N -> S (IN STRAIN CCRC 21164).
FT VARIANT 350 350 D -> A (IN STRAINS CCRC 21164 AND ATCC
FT VARIANT 479 479 26077).
FT VARIANT 479 479 L -> S (IN STRAINS CCRC 21164 AND ATCC
FT VARIANT 483 483 26077).
FT VARIANT 512 AA; 56527 MM; 857552BCF60F965 CRC64;
SO SEQUENCE

```

```

Query Match 64.18; Score 50; DB 1; Length 512;
Best Local Similarity 61.58; Pred. No. 0.61;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1 AYHGYWOODIYSL 13
    ||||| :| |
Db 114 AYHGYWAKNIDEL 126

```

```

RESULT 11
AMY_BACCI STANDARD; PRT; 528 AA.
ID AMY_BACCI
AC P08137;
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amyglase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
DE Bacillus circulans.
OS Bacillus; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
OX NCBI_TaxID=1397;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87246076; PubMed=3109866;
RX Nishizawa M., Ozawa F., Hishinuma F.;
RT "Molecular cloning of an amyglase gene of Bacillus circulans.";
RL DNA 6:255-265(1987).
CC -1 CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M16657; AAA22229.1; -
DR PIR: A29083; ALBSU.
DR HSSP: P26827; LCIU.
DR InterPro: IPR000461; Alpha-amyglase.
DR Pfam: PF00128; alpha-amyglase_C; 1.
KW Hydroglase; Glycosylase; Carbohydrate metabolism; Signal.
FT SIGNAL 1 28
FT CHAIN 29 528 ALPHA-AMYLASE.
FT ACT_SITE 258 258 BY SIMILARITY.
FT ACT_SITE 262 262 BY SIMILARITY.
FT ACT_SITE 357 357 BY SIMILARITY.
SQ SEQUENCE 528 AA; 57940 MW; 455740202FA29753 CRC64;

```

```

Query Match 62.88; Score 49; DB 1; Length 528;
Best Local Similarity 77.88; Pred. No. 0.92;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 AYHGYWOOD 9
    ||||| :|

```

```

Db 125 AYHGYWARD 133
RESULT 12
CDGT_BACS3 STANDARD; PRT; 712 AA.
ID CDGT_BACS3
AC P09121;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Cyclomaltoextrin glucanotransferase precursor (EC 2.4.1.19)
DE (Cyclodextrin-glycosyltransferase) (Cgtase).
GN Cgt.
OS Bacillus sp. (strain 38-2).
OC Bacillus; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
OX NCBI_TaxID=1412;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-44.
RX MEDLINE=89036108; PubMed=2972812;
RA Kaneko T., Hamamoto T., Horikoshi K.;
RT "Molecular cloning and nucleotide sequence of the cyclomaltoextrin
RT glucanotransferase gene from the alkalophilic Bacillus sp. strain no.
RT 38-2.";
RL J. Gen. Microbiol. 134:97-105(1988).
RN [2]
RP SEQUENCE OF 1-586 FROM N.A.
RA Hamamoto T., Kaneko T., Horikoshi K.;
RT "Nucleotide sequence of the cyclomaltoextrin glucanotransferase
RT (Cgtase) gene from alkalophilic Bacillus sp. strain No. 38-2.";
RL Agric. Biol. Chem. 51:2019-2022(1987).
CC -1 CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation
CC of a 1,4-alpha-D-glucosidic bond.
CC -1 COPFACTOR: BINDS TWO CALCIUM IONS.
CC -1 SUBUNIT: MONOMER.
CC -1 MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE
CC IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND
CC IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER
CC ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN
CC ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE
CC MALTOOLIGOSACCHARIDE PRODUCED.
CC -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M19880; AAA22309.1; -
DR EMBL: D00129; BAA00077.1; -
DR PIR: S24193; ALBSG3.
DR HSSP: P05618; IPRAM.
DR InterPro: IPR000461; Alpha-amyglase.
DR InterPro: IPR002044; CBD_4.
DR InterPro: IPR002909; IPT_TIG.
DR Pfam: PF00128; alpha-amyglase; 1.
DR Pfam: PF00806; alpha-amyglase_C; 1.
DR Pfam: PF00686; CBD_4; 1.
DR Pfam: PF01833; TIG_1.
DR ProDom: PD001568; CBD_4; 1.
KW Transferase; Glycosyltransferase; Calcium; Signal.
FT SIGNAL 1 27
FT CHAIN 28 712 CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
FT DOMAIN 28 165 A1.
FT DOMAIN 166 229 B.
FT DOMAIN 230 433 A2.
FT DOMAIN 434 522 C.
FT DOMAIN 523 608 D.
FT DOMAIN 609 712 E.

```

FT DISUFLD 70 77 BY SIMILARITY.
 FT ACT_SITE 256 256 BY SIMILARITY.
 FT ACT_SITE 284 284 BY SIMILARITY.
 FT ACT_SITE 355 355 BY SIMILARITY.
 FT ACT_SITE 582 586 VPGI -> SWRL (IN REF. 2).
 FT CONFLICT 582 586
 SO SEQUENCE 712 AA: 78249 MW: 4FAABF70BEF81BF9 CRC64;

Query Match 62.8%; Score 49; DB 1; Length 712;
 Best Local Similarity 77.8%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGYWOOD 9
 DB 123 AYHGYWARD 131

RESULT 13
 AMYR_BACS8 STANDARD; PRT: 713 AA.
 AC P17692.
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Raw-starch-digesting amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase)
 OS Bacillus sp. (Strain B1018).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/staphylococcus group; Bacillus.
 OX NCBI_TaxID=1417;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-47.
 RX MEDLINE=90147765; PubMed=1689153;
 RA Riktor P., Tsuchagoshi N., Uda S.;
 RT "Nucleotide sequence of the raw-starch-digesting amylase gene from RT Bacillus sp. B1018 and its strong homology to the cyclodextrin RT glucanotransferase genes."
 RL Biochem. Biophys. Res. Commun. 166:630-636(1990).
 CC -1- FUNCTION: THIS ENDO-TYPE ADSORBABLE AMYLASE IS CAPABLE TO CC DIEST RAW-STARCH.
 CC -1- CATALYTIC ACTIVITY: Endohydrolays of 1,4-alpha-glucosidic CC linkages in oligosaccharides and polysaccharides
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M33302; AAA2239.1; -;
 DR EMBL: D90112; BAA14140.1; -;
 DR PIR: A34648; A34648.
 DR PIR: S09196; S09196.
 DR HSSP: P43379; ICDG.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amylase_1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR Prodom: PD001568; CBD_4; 1.
 KM Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 713 RAW-STARCH-DIGESTING AMYLASE.
 FT ACT_SITE 256 256 BY SIMILARITY.
 FT ACT_SITE 260 260 BY SIMILARITY.

FT ACT_SITE 355 355 BY SIMILARITY.
 SO SEQUENCE 713 AA: 77420 MW: 85FB616DA687B888 CRC64;

Query Match 62.8%; Score 49; DB 1; Length 713;
 Best Local Similarity 77.8%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGYWOOD 9
 DB 123 AYHGYWARD 131

RESULT 14
 CDG2_PAEMA STANDARD; PRT: 713 AA.
 AC P31835.
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Cyclomalto-dextrin glucanotransferase precursor (EC 2.4.1.19)
 DE (Cyclodextrin-glycosyltransferase) (CGTase).
 OS Paenibacillus macerans (Bacillus macerans).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/staphylococcus group; Paenibacillus.
 OX NCBI_TaxID=44252;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-37.
 RA Sugimoto T., Kubota M., Sakai S.;
 RT "Polypeptide possessing cyclomalto-dextrin glucanotransferase RT activity."
 RL Patent number GB2169902, 23-JUL-1986.
 CC -1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation CC of a 1,4-alpha-D-glucosidic bond.
 CC -1- COFACTOR: BINDS TWO CALCIUM IONS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS. THE ONE CC IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GUCCOSIDIC BOND CC IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER CC ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN CC ALPHA-1,4-GUCCOSIDIC LINKAGE FOR CYCLIZING THE CC MATOOLIGOSACCHARIDE PRODUCED.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 DR PIR: S26589; ALEBXR.
 DR HSSP: P43379; ICDG.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amylase_1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR Prodom: PD001568; CBD_4; 1.
 KM Transferase; Glycosyltransferase; Calcium; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 713 CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
 FT DOMAIN 166 165 A1.
 FT DOMAIN 166 229 B.
 FT DOMAIN 230 434 A2.
 FT DOMAIN 435 522 C.
 FT DOMAIN 523 609 D.
 FT DOMAIN 610 713 E.
 FT ACT_SITE 256 256 BY SIMILARITY.
 FT ACT_SITE 285 285 BY SIMILARITY.
 FT ACT_SITE 356 356 BY SIMILARITY.
 SO SEQUENCE 713 AA: 76857 MW: 5A287BCC4AAFE635 CRC64;

Query Match 62.8%; Score 49; DB 1; Length 713;
 Best Local Similarity 77.8%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGYWOOD 9
 Db 123 AYHGYWPRD 131

RESULT 15

CGGT_BACSO
 ID CDGT_BACSO STANDARD; PRT: 713 AA.
 AC P05618;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Cyclomaltodextrin glucanotransferase precursor (EC 2.4.1.19)
 DE (Cyclodextrin-glycosyltransferase) (Cgtase).
 GN Cgt.
 OS Bacillus sp. (strain 1011).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87308036; PubMed-2957361;
 RA Kimura K., Kataoka S., Ishii Y., Takano T., Yamane K.;
 RT "Nucleotide sequence of the beta-cyclodextrin glucanotransferase gene
 RT of alkalophilic Bacillus sp. strain 1011 and similarity of its amino
 RT acid sequence to those of alpha-amylases.";
 RL J. Bacteriol. 169:4399-4402(1987).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RA Harata K., Haga K., Nakamura A., Aoyagi M., Yamane K.;
 RT "X-ray structure of cyclodextrin glucanotransferase from alkalophilic
 RT Bacillus Sp. 1011. Comparison of two independent molecules at 1.8-A
 RT resolution.";
 RL Acta Crystallogr. D 52:1136-1145(1996).
 CC -1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation
 CC of a 1,4-alpha-D-glucosidic bond.
 CC -1- COFACTOR: BINDS TWO CALCIUM IONS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE
 CC IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND
 CC IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER
 CC ACTIVITIES, INCLUDING THE RECONSTRUCTION OF AN
 CC ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE
 CC MALTOOLIGOSACCHARIDE PRODUCED.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17366; AAA2308.1; -.
 DR PIR: A26678; ALBSG1.
 DR PDB: 1PAM; 1JAN-97.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR ProDom: PD001568; CBD_4; 1.
 DR Transferrase; Glycosyltransferase; Calcium; Signal; 3D-structure.
 KW TRANSFERRASE; GLYCOSYLTRANSFERASE;
 FT CHAIN 1 27
 FT SIGNAL 1 27
 FT CHAIN 28 713 CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
 FT DOMAIN 28 165 A1.
 FT DOMAIN 166 229 B.
 FT DOMAIN 230 433 A2.
 FT DOMAIN 434 522 C.

FT DOMAIN 523 609 D.
 FT DOMAIN 610 713 E.
 FT DISULFD 70 77 BY SIMILARITY.
 FT ACT_SITE 256 256 BY SIMILARITY.
 FT ACT_SITE 284 284 BY SIMILARITY.
 FT ACT_SITE 335 335 BY SIMILARITY.
 SQ SEQUENCE 713 AA; 78340 MW; 524B259526B56C52 CRC64;

Query Match 62.8%; Score 49; DB 1; Length 713;
 Best Local Similarity 77.8%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 AYHGYWOOD 9
 Db 123 AYHGYWARD 131

Search completed: June 13, 2002, 09:11:49
 Job time: 1029 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 13, 2002, 08:55:49 ; Search time 73.31 Seconds
(without alignments)
17,039 Million cell updates/sec

Title: US-09-710-339-2_COPY_98_110
Perfect score: 78
Sequence: 1 AVHGYWQDDIVSL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR-71:*
1: piri:*
2: piri:*
3: piri:*
4: piri:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	78	100.0	498	2 A48305	alpha-amylase (EC
2	78	100.0	499	1 A48305	alpha-amylase (EC
3	78	100.0	499	1 A48305	alpha-amylase (EC
4	78	100.0	499	2 B48305	alpha-amylase (EC
5	78	100.0	499	2 J50863	alpha-amylase (EC
6	78	100.0	499	2 J50863	alpha-amylase (EC
7	78	100.0	499	2 J50863	alpha-amylase (EC
8	78	100.0	499	2 J50863	alpha-amylase (EC
9	78	100.0	499	2 J50863	alpha-amylase (EC
10	78	100.0	499	2 J50863	alpha-amylase (EC
11	78	100.0	499	2 J50863	alpha-amylase (EC
12	78	100.0	499	2 J50863	alpha-amylase (EC
13	78	100.0	499	2 J50863	alpha-amylase (EC
14	78	100.0	499	2 J50863	alpha-amylase (EC
15	78	100.0	499	2 J50863	alpha-amylase (EC
16	78	100.0	499	2 J50863	alpha-amylase (EC
17	78	100.0	499	2 J50863	alpha-amylase (EC
18	78	100.0	499	2 J50863	alpha-amylase (EC
19	78	100.0	499	2 J50863	alpha-amylase (EC
20	78	100.0	499	2 J50863	alpha-amylase (EC
21	78	100.0	499	2 J50863	alpha-amylase (EC
22	78	100.0	499	2 J50863	alpha-amylase (EC
23	78	100.0	499	2 J50863	alpha-amylase (EC
24	78	100.0	499	2 J50863	alpha-amylase (EC
25	78	100.0	499	2 J50863	alpha-amylase (EC
26	78	100.0	499	2 J50863	alpha-amylase (EC
27	78	100.0	499	2 J50863	alpha-amylase (EC
28	78	100.0	499	2 J50863	alpha-amylase (EC
29	78	100.0	499	2 J50863	alpha-amylase (EC

30	49	62.8	713	2 A58800	cyclomaltoextrin
31	49	62.8	718	1 ALBSMX	cyclomaltoextrin
32	49	62.8	718	1 ALBSG6	cyclomaltoextrin
33	49	62.8	718	1 ALBSG6	cyclomaltoextrin
34	49	62.8	718	1 ALBSG6	cyclomaltoextrin
35	49	62.8	718	1 ALBSG6	cyclomaltoextrin
36	49	62.8	718	1 ALBSG6	cyclomaltoextrin
37	49	62.8	718	1 ALBSG6	cyclomaltoextrin
38	49	62.8	718	1 ALBSG6	cyclomaltoextrin
39	49	62.8	718	1 ALBSG6	cyclomaltoextrin
40	49	62.8	718	1 ALBSG6	cyclomaltoextrin
41	49	62.8	718	1 ALBSG6	cyclomaltoextrin
42	49	62.8	718	1 ALBSG6	cyclomaltoextrin
43	49	62.8	718	1 ALBSG6	cyclomaltoextrin
44	49	62.8	718	1 ALBSG6	cyclomaltoextrin
45	49	62.8	718	1 ALBSG6	cyclomaltoextrin

ALIGNMENTS

RESULT 1
A48305
alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: A48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J
Curr. Genet. 17, 203-212, 1990
A:title: Cloning, characterization, and expression of two alpha-amylase genes from As
A:Reference number: A48305; MUID:90254827
A:Accession: A48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-498 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMT>

Query Match 100.0%; Score 78; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AVHGYWQDDIVSL 13
DB 99 AVHGYWQDDIVSL 111
RESULT 2
A48305
alpha-amylase (EC 3.2.1.1) A precursor (validated) - Aspergillus oryzae
N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2000
C:Accession: S04548; A33214; J50240; A91930; A93767; A10627
R:Wise, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989
A:title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
A:Reference number: S04548; MUID:89237897
A:Accession: S04548
A:Molecule type: DNA
A:Residues: 1-499 <WIR>
A:Cross-references: EMBL:X12725; NID:92430; PIDN:CAA31218.1; PID:9295921
A:Accession: A33214
A:Molecule type: mRNA
A:Residues: 1-499 <WIR>
A:Cross-references: GB:X12725; NID:92430; PIDN:CAA31218.1; PID:9295921
R:Genes, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amyase genes, each containing
 A:Reference number: JS0240; MUID:89378767
 A:Accession: JS0240
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A:Genetics: AMY2
 A>Note: the authors refer to this as isozyme II
 R:Isomura, S.; Ikenaka, T.
 J. Biochem. 74, 1-10, 1973
 A:Reference number: A91930; MUID:74001521
 A:Accession: A91930
 A:Molecule type: protein
 A:Residues: 206-225 <ISE>
 R:Narita, K.
 Proc. Jpn. Acad. 51, 285-290, 1975
 A:Reference number: A93767
 A:Accession: A93767
 A:Molecule type: protein
 A:Residues: 434-443,446-447, 'Q',449-458, 'GTTV',459-464,467-468, 'B',470, 'B',472-499 <NAR>
 R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amyase A.
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms
 A:Reference number: A37454; MUID:84212370
 A:Accession: A37454
 R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.
 submitted to the Brookhaven Protein Data Bank, August 1992
 A:Reference number: A51348; PDB:61HA
 A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497
 A:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: <AMY1>
 A:Gene: amy1
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Genetics: <AMY2>
 A:Gene: amy2; Amy11
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amyase; alpha-amyase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; polys
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:122-499/Product: alpha-amyase 1 #status experimental <MAT>
 F:194-321/Domain: alpha-amyase core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 78; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 |||||||||||||
 DB 99 AYHGWOODIYSL 111

RESULT 3
 ALAS3
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 N:Alternate names: alpha-amyase isozyme I; glycogenase; Taka-amyase A
 C:Species: Aspergillus oryzae
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713
 R:Witresl, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A:Title: Three alpha-amyase genes of Aspergillus oryzae exhibit identical intron-exon
 A:Reference number: S04548; MUID:89237897
 A:Accession: S04549
 A:Molecule type: DNA
 A:Residues: 1-499 <WIR>

A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 A:Accession: A33215
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-499 <WIR>
 A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 R:Genes, M.J.; Dove, M.D.; Seligy, V.L.
 Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amyase genes, each contain
 A:Reference number: JS0240; MUID:89378767
 A:Accession: A44713
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A>Note: the authors refer to this as isozyme I
 R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amyase A.
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms
 A:Reference number: A37454; MUID:84212370
 A:Accession: A37454
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics:
 A:Gene: amy1
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amyase; alpha-amyase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:122-499/Product: alpha-amyase 3 #status experimental <MAT>
 F:194-321/Domain: alpha-amyase core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 78; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 |||||||||||||
 DB 99 AYHGWOODIYSL 111

RESULT 4
 B48305
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 C:Species: Aspergillus awamori
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
 C:Accession: B48305
 R:Kotman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J
 Curr. Genet. 17, 203-212, 1990
 A:Title: Cloning, characterization, and expression of two alpha-amyase genes from As
 A:Reference number: A48305; MUID:90254827
 A:Accession: B48305
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499 <KOR>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amyase; alpha-amyase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amyase core homology <AMY>

Query Match 100.0%; Score 78; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13

Db 99 AYHGWOODIYSL 111

RESULT 5

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
 C:Species: Aspergillus sp.
 C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
 C:Accession: J50663
 R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
 Biosci. Biotechnol. Biochem. 56, 174-179, 1992.
 A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirossamii and its expression
 A:Reference number: J50663; MUID:92323146
 A:Accession: J50663
 A:Molecule type: mRNA
 A:Residues: 1-499 <SHI>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domains: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <ALP>
 F:194-321/Domains: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 78; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
 Db 99 AYHGWOODIYSL 111

RESULT 6

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 N:Alternate names: glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
 C:Accession: J70466
 R:Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
 Agric. Biol. Chem. 53, 593-599, 1989
 A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus
 A:Reference number: J70466
 A:Accession: J70466
 A:Molecule type: DNA
 A:Residues: 1-499 <TAD>
 C:Comment: See also PIR:JK0201 and PIR:JS0240.
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics:
 A:Insertions: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domains: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <AMY>
 F:194-321/Domains: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 78; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
 Db 99 AYHGWOODIYSL 111

Db 99 AYHGWOODIYSL 111

RESULT 7

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 N:Alternate names: Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
 C:Accession: JN0588
 R:Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kiritla, N.; Tsuboi, A.; Udaoka, S.
 Gene 84, 319-327, 1989
 A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
 A:Reference number: JN0588; MUID:90128276
 A:Accession: JN0588
 A:Molecule type: mRNA
 A:Residues: 1-499 <TSU>
 C:Comment: The alpha amylases are encoded by multigene family.
 C:Genetics:
 A:Gene: Taa-G1
 A:Insertions: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domains: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <AMY>
 F:194-321/Domains: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 78; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
 Db 99 AYHGWOODIYSL 111

RESULT 8

alpha-amylase a precursor - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
 C:Accession: T38770
 R:Skellton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: 221751
 A:Accession: T38770
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-513 <SKE>
 A:Cross-references: EMBL:298762; PIDN:CAB11471.1; GSPDB:GN00066; SPDB:SPAC4A8.01
 A:Experimental source: strain 972h-; cosmid c4A8
 C:Genetics:
 A:Gene: SPDB:SPAC4A8.01
 A:Map position: 1
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

Query Match 85.9%; Score 67; DB 2; Length 513;
 Best Local Similarity 76.9%; Pred. No. 0.0024;
 Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
 Db 100 AYHGWOODIYSL 112

RESULT 9

S72270

alpha-amylase (EC 3.2.1.1) precursor - *Cryptococcus* sp. (strain CS2)
 C:Species: *Cryptococcus* sp.
 A:Variety: strain CS2
 C>Date: 23-Apr-1998 #sequence_revision 08-May-1998 #text_change 20-Jun-2000
 C:Accession: S72270
 R:RefSeq: H.; Chino, M.; Kato, M.; Iimura, Y.
 Biochem. J. 318, 989-996, 1996
 A>Title: Raw-starch-digesting and thermostable alpha-amylase from the yeast *Cryptococcus*
 A:Reference number: S72270; MUID:96433120
 A:Accession: S72270
 A:Molecule type: DNA
 A:Residues: 1-631 <IE>
 A:Cross-references: EMBL:D83540; NID:g1595852; PIDN:BA12010.1; PID:g1595853
 A:Experimental source: strain S-2
 C:Genetics:
 A:Gene: amy-CS2
 A:Introns: 289/2; 326/2
 C:Function:
 A:Pathway: glycogen/starch degradation
 C:Superfamily: *Aspergillus* alpha-amylase; alpha-amylase core homology
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-631/Product: alpha-amylase #status predicted <MAT>
 F:206-335/Domain: alpha-amylase core homology <AMY>

Query Match 84.6%; Score 66; DB 2; Length 631;
 Best Local Similarity 76.9%; Pred. No. 0.0044;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||| |||||

DB 105 AYHGWOODIYEL 117

RESULT 10

alpha-amylase (EC 3.2.1.1) - *Aspergillus niger*
 C:Species: *Aspergillus niger*
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A35282
 R:Boel, E.; Brady, L.; Brzozowski, A.M.; Derewenda, Z.; Dodson, G.G.; Jensen, V.J.; Pete
 Biochemistry 29, 6244-6249, 1990
 A>Title: Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-angstrom
 A:Reference number: A35282; MUID:91002514
 A:Accession: A35282
 A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA
 A:Residues: 1-484 <BO>
 C:Function:
 A>Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: *Aspergillus* alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolyase; polysaccharide degradation
 F:173-300/Domain: alpha-amylase core homology <AMY>

Query Match 82.1%; Score 64; DB 1; Length 484;
 Best Local Similarity 76.9%; Pred. No. 0.0071;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||| |||||

DB 78 AYHGWOODIYDV 90

RESULT 11

alpha-amylase (EC 3.2.1.1) precursor - yeast (*Schwannomyces occidentalis*)
 N:Alternate names: alpha-1,4 glucanohydrolyase
 C:Species: *Schwannomyces occidentalis*
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S33921

R:Claros, M.G.; Abarca, D.; Fernandez-Lobato, M.; Jimenez, A.
 Curr. Genet. 24, 75-83, 1993
 A>Title: Molecular structure of the SMA2 gene encoding an AMY1-related alpha-amylase
 A:Reference number: S33921; MUID:93365041
 A:Accession: S33921
 A:Molecule type: DNA
 A:Residues: 1-507 <CL>
 A:Cross-references: EMBL:X73497; NID:g396561; PIDN:CA51912.1; PID:g396562
 C:Genetics:
 A:Gene: SMA2
 C:Function:
 A>Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: *Aspergillus* alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolyase; polysaccharide degradation
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-507/Product: alpha-amylase #status predicted <MAT>
 F:205-332/Domain: alpha-amylase core homology <AMY>
 F:134,229/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.1%; Score 64; DB 2; Length 507;
 Best Local Similarity 69.2%; Pred. No. 0.0075;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||| |||||

DB 110 AYHGWOODIYAI 122

RESULT 12

alpha-amylase (EC 3.2.1.1) - *Aspergillus oryzae*
 N:Alternate names: glycogenase; Taka-amylase A
 C:Species: *Aspergillus oryzae*
 C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
 C:Accession: JK0201
 R:Roda, H.; Kondo, K.; Narita, K.
 Proc. Jpn. Acad. 58B, 208-212, 1982
 A>Title: The complete amino acid sequence of Taka-amylase A.
 A:Reference number: JK0201
 A:Accession: JK0201
 A:Molecule type: protein
 A:Residues: 1-478 <OD>
 C:Comment: One atom of calcium per molecule is essential for the activity.
 C:Comment: This enzyme is a glycoprotein.
 C:Comment: See also PIR:JT0466 and PIR:J50240.
 C:Function:
 A>Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: *Aspergillus* alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolyase; polysaccharide degradation
 F:173-300/Domain: alpha-amylase core homology <AMY>
 F:197/binding site: carbohydrate (Asn) (covalent) #status experimental
 F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 79.5%; Score 62; DB 2; Length 478;
 Best Local Similarity 84.6%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||| |||||

DB 78 AYTGWQTDIYSL 90

RESULT 13

alpha-amylase (EC 3.2.1.1) precursor - yeast (*Lipomyces kononenkoae*)
 N:Alternate names: LKA1 protein; raw starch-degrading amylase
 N:Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)
 C:Species: *Lipomyces kononenkoae*
 C>Date: 07-Feb-1996 #sequence_revision 23-Aug-1996 #text_change 18-Jun-1999

C:Accession: J04510; PC4116
 R:Steyn, A.J.C.; Marmur, J.; Pretorius, I.S.
 Gene 166, 65-71, 1995
 A:Title: Cloning, sequence analysis and expression in yeasts of a cDNA containing a lipC
 A:Reference number: J04510; MUID:96105202
 A:Accession: J04510
 A:Molecule type: mRNA
 A:Residues: 1-624 <STE>
 A:Cross-references: GB:U30376; NID:g1173536; PIDN:AAC49622.1; PID:g1173537
 A:Experimental source: strain IGC4052B
 A:Accession: PC4116
 A:Molecule type: protein
 A:Residues: 29-44 <ST>
 A:Experimental source: IGC4052B
 C:Genetics:
 A:Gene: LKAL
 C:Function:
 A:Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages
 A:Pathway: glycogen/starch degradation
 C:Superfamily: lipomyces alpha-amylase; alpha-amylase core homology; glucoamylase starch
 C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-624/Product: alpha-amylase #status predicted <MAT>
 F:48-141/Domain: glucoamylase starch-binding domain homology <MAT>
 F:120-447/Domain: alpha-amylase core homology <AM>
 F:177-185,297-311,367-430,587-622/Disulfide bonds: #status predicted
 F:304,344/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:357,377,444/Active site: His, Glu, Asp #status predicted

Query Match 79.5%; Score 62; DB 1; Length 624;
 Best Local Similarity 69.2%; Pred. No. 0.02;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AYHGWOODIYSL 13
 ||||| :||:|
 DB 225 AYHGWMKDIKFL 237

RESULT 14
 S31478
 alpha-amylase (EC 3.2.1.1) - Thermactinomyces vulgaris
 C:Species: Thermactinomyces vulgaris
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 22-Oct-1999
 C:Accession: S31478
 R:Hoefmeester, B.; Koenig, S.; Hoang, V.; Engel, J.; Mayer, G.; Heese, O.; Hansen, G.; H
 submitted to the EMBL Data Library, December 1992
 A:Description: The gene amyTV coding for a non-gluco-genic alpha-amylase from Thermactin
 A:Reference number: S31478
 A:Accession: S31478
 A:Molecule type: DNA
 A:Residues: 1-482 <HO>
 A:Cross-references: EMBL:X69807; NID:g48289; PIDN:CAA49465.1; PID:g48290
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 F:194-319/Domain: alpha-amylase core homology <AMY>

Query Match 75.6%; Score 59; DB 2; Length 482;
 Best Local Similarity 69.2%; Pred. No. 0.048;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
 ||||| :||:|
 DB 107 AYHGWTYDFYSV 119

RESULT 15
 ALBYAF
 alpha-amylase (EC 3.2.1.1) precursor - yeast (Saccharomyces fibuligera)

C:Species: Saccharomyces fibuligera
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 18-Jun-1999
 C:Accession: S00064
 R:Itou, T.; Yamashita, I.; Fukui, S.
 FEBS Lett. 219, 339-342, 1987
 A:Title: Nucleotide sequence of the alpha-amylase gene (ALP1) in the yeast Saccharomy
 A:Reference number: S00064; MUID:87276512
 A:Accession: S00064
 A:Molecule type: DNA
 A:Residues: 1-494 <IT>
 A:Cross-references: EMBL:X05791; NID:g4847; PIDN:CAA29233.1; PID:g4848
 A>Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in having 366-
 C:Genetics:
 A:Gene: ALP1
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; extracellular protein; glycoprotein; glycosidase; hydrolase; met
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-494/Product: alpha-amylase #status predicted <MAT>
 F:200-327/Domain: alpha-amylase core homology <AM>
 F:57-65,177-191,267-310,462-493/Disulfide bonds: #status predicted
 F:148,189,202,237/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:224/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:233,257,324/Active site: Asp, Glu, Asp #status predicted

Query Match 74.4%; Score 58; DB 1; Length 494;
 Best Local Similarity 61.5%; Pred. No. 0.072;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AYHGWOODIYSL 13
 ||||| :||:|
 DB 105 AYHGWMKNIYKI 117

Search completed: June 13, 2002, 08:55:49
 Job time: 369 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 13, 2002, 08:53:25 ; Search time 158.68 seconds
(without alignments)
9.100 Million cell updates/sec

Title: US-09-710-339-2_COPY_98_110

Perfect score: 78

Sequence: 1 AHGWTQDDIYSL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	78	100.0	478	15	AAR46065	Mutant alpha-amylase
2	78	100.0	478	16	AAR72450	Aspergillus oryzae
3	78	100.0	478	17	AAR78270	Aspergillus oryzae
4	78	100.0	478	17	AAW14500	Aspergillus oryzae
5	78	100.0	498	22	AAH84206	Amino acid sequenc
6	69	88.5	55	17	AAR88213	Alpha-amylase frag
7	69	88.5	493	17	AAR88212	Alpha-amylase. Th
8	62	79.5	478	16	AAR79025	Mature taka-amylas
9	58	74.4	468	13	AAR24136	Alpha-amylase vari
10	58	74.4	494	8	AAW70571	Alpha-amylase gene
11	54	69.2	468	15	AAR63184	Variant alpha amyl

12	53	67.9	468	15	AAR63185	Variant alpha amyl
13	50	64.1	452	15	AAR63186	Variant alpha amyl
14	50	64.1	511	9	AAW1161	Recombinant alpha-
15	50	64.1	511	9	AAW1160	Sequence of alpha-
16	50	64.1	512	11	AAR07574	Alpha-amylase enco
17	49	62.8	468	15	AAR63187	Variant alpha amyl
18	49	62.8	656	11	AAR06109	Sequence of cyclom
19	49	62.8	666	7	AAW61518	Sequence of cyclom
20	49	62.8	666	17	AAW06773	Wild type CGTase.
21	49	62.8	666	17	AAW17592	Bacillus CGTase va
22	49	62.8	666	17	AAW17593	Bacillus CGTase va
23	49	62.8	666	17	AAW17594	Bacillus CGTase va
24	49	62.8	666	17	AAW17595	Bacillus CGTase va
25	49	62.8	666	17	AAW17596	Bacillus CGTase va
26	49	62.8	666	17	AAW17597	Bacillus CGTase va
27	49	62.8	666	17	AAW17598	Bacillus CGTase va
28	49	62.8	666	17	AAW17599	Bacillus CGTase va
29	49	62.8	666	17	AAW17600	Bacillus CGTase va
30	49	62.8	666	17	AAW17602	Bacillus CGTase va
31	49	62.8	666	17	AAW17603	Bacillus CGTase va
32	49	62.8	666	17	AAW17604	Bacillus CGTase va
33	49	62.8	666	17	AAW17605	Bacillus CGTase va
34	49	62.8	666	20	AAW02599	A cyclomalitodextri
35	49	62.8	687	17	AAW17601	Bacillus CGTase va
36	49	62.8	688	17	AAW17568	Bacillus CGTase va
37	49	62.8	688	17	AAW17572	Bacillus CGTase va
38	49	62.8	688	17	AAW17576	Bacillus CGTase va
39	49	62.8	712	12	AAW10051	Cyclomalitodextrin
40	49	62.8	713	11	AAR06110	Sequence of cyclom
41	49	62.8	713	12	AAR10052	Cyclomalitodextrin
42	49	62.8	713	12	AAR12743	Enzyme with starch
43	49	60.3	655	13	AAR20139	Sequence encoded b
44	47	60.3	715	13	AAR20138	Sequence encoded b
45	46	59.0	170	22	AAW25600	Human G Protein-Co

ALIGNMENTS

RESULT	1
AAW46065	standard; protein; 478 AA.
XX	
AC	AAW46065;
XX	
DT	18-JUL-1994 (first entry)
XX	
DE	Mutant alpha-amylase.
XX	
KM	Methionine substitution; stability; activity; detergent;
KW	dishwashing agents; liquifaction agents.
XX	
OS	Aspergillus oryzae.
XX	
PN	WO9402597-A.
XX	
PD	03-FEB-1994.
XX	
PF	06-JUL-1993; 93WO-DK00230.
XX	
PR	23-JUL-1992; 92DK-0000946.
PR	16-DEC-1992; 92DK-0001503.
PR	15-MAR-1993; 93DK-0000292.
XX	
PA	(NOVO) NOVO-NORDISK AS.
XX	
PI	Bisgard-frantzen H, Svendsen A;
XX	
DR	WPI; 1994-048855/06.
XX	
PT	Mutant alpha-amylase from Bacillus species comprising a
PT	methionine substitution - with improved stability and activity at
PT	low pH, for use in detergents, dishwashing agents and

PT Liquefaction agents
 XX
 PS Claim 1; Page 7; 20pp; English.
 CC The sequence os that of the Aspergillus oryzae alpha amylase, sold
 CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
 CC be mutated by substitution of one or more of its methionine residues
 CC for any amino acid other than cysteine. The mutant alpha-amylase
 CC exhibits a better activity level and better stability in the
 CC presence of oxidising agents than previous mutant alpha amylases,
 CC and improved thermostability at moderately low pH. The enzyme can
 CC be used as an additive for detergents, dishwashing agents and
 CC liquifaction agents.
 XX
 SQ Sequence 478 AA;

Query Match 100.0%; Score 78; DB 15; Length 478;
 Best Local Similarity 100.0%; Pred. NO. 0.00036;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AYHGWMQODIYSL 13
 ||||||||||||
 Db 78 ayhgywqgdyls1 90

RESULT 2

AA872450
 ID AA872450 standard; Protein; 478 AA.
 AC AA872450;
 XX
 DT 01-DEC-1995 (first entry)
 XX
 DE Aspergillus oryzae alpha amylase (mature protein).
 KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
 KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
 KW Bacillus steaerothermophilus; dyeing; bleaching; scouring; textile;
 KW thermostable.
 XX
 OS Aspergillus oryzae.
 XX
 PN W09510603-A.
 XX
 PD 20-APR-1995.
 XX
 PF 05-OCT-1994; 94WO-DK00370.
 XX
 PR 08-OCT-1993; 93DK-0001133.
 PR 02-FEB-1994; 94DK-0000140.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Svendsen A, Thelersen M;
 PI Van der zee P, Bisgardtrantzen H, Borchert T;
 DR WPI; 1995-161790/21.
 XX
 PT New Bacillus derived alpha-amylase variants - having amino acid
 PT modifications to improve washing and/or dishwashing performance
 XX
 PS Disclosure; Page 75-76; 105pp; English.
 CC Variant alpha amylase enzymes which have improved washing and/or
 CC as detergent additives. The enzymes have one or more amino acid
 CC residues added, deleted or substituted. The variants can also be
 CC used for textile desizing prior to scouring, bleaching and dyeing.
 CC The variants have improved thermostability, acid/alkaline stability;
 CC low temperature optimum; pH optimum; higher hydrolysis velocity and
 CC improved tolerance to other composition constituents, e.g. oxidation
 CC agents.
 XX

SQ Sequence 478 AA;

Query Match 100.0%; Score 78; DB 16; Length 478;
 Best Local Similarity 100.0%; Pred. NO. 0.00036;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AYHGWMQODIYSL 13
 ||||||||||||
 Db 78 ayhgywqgdyls1 90

RESULT 3

AA878270
 ID AA878270 standard; Protein; 478 AA.
 AC AA878270;
 XX
 DT 17-JAN-1996 (first entry)
 XX
 DE Aspergillus oryzae alpha amylase (mature protein).
 KW Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
 KW starch; thermostable; methionine; Bacillus licheniformis;
 KW Bacillus amyloliquefaciens; Bacillus steaerothermophilus;
 KW Aspergillus oryzae.
 XX
 OS Aspergillus oryzae.
 XX
 PN W09521247-A1.
 XX
 PD 10-AUG-1995.
 XX
 PF 05-OCT-1994; 94WO-DK00371.
 XX
 PR 02-FEB-1994; 94DK-0000141.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Marcher D, Nilsson TE, Pedersen HH, Toft AH;
 DR WPI; 1995-28367/37.
 XX
 PT Use of an oxidation stable alpha-amylase - for simultaneous desizing
 PT and bleaching or scouring of fabrics contg. starch or starch derivs.
 XX
 PS Disclosure; Page 25-26; 37pp; English.
 CC Oxidation stable alpha amylases can be used for the simultaneous
 CC desizing and bleaching or scouring of a fabric comprising starch or
 CC starch derivatives. They exhibit a better heat stability.
 CC especially in the presence of oxidising agents. They are obtained
 CC from a parent alpha amylase by replacing one or more methionine
 CC residues with any amino acid different from Cys or Met, preferably
 CC Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is
 CC pref. derived from a Bacillus species, although alpha amylases of
 CC fungal origin can also be used. This sequence is the wild type
 CC (unmodified) alpha amylase of Aspergillus oryzae.
 XX
 SQ Sequence 478 AA;

Query Match 100.0%; Score 78; DB 16; Length 478;
 Best Local Similarity 100.0%; Pred. NO. 0.00036;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWMQODIYSL 13
 ||||||||||||
 Db 78 ayhgywqgdyls1 90

RESULT 4

AA814500

ID AAW14500 standard; protein; 478 AA.
 XX AAW14500;
 AC
 XX
 DT 04-JUN-1997 (first entry)
 XX
 DE Aspergillus oryzae alpha-amylase (mature protein).
 XX
 KW alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;
 KW Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;
 KW calcium dependency; substrate binding; stability; pH optimum;
 KW thermostability; cleavage; oligosaccharide substrate; dishwashing;
 KW washing; detergent additive; fabric desizing; starch liquefaction;
 KW sweetener; ethanol production; variant.
 XX
 OS Aspergillus oryzae.
 XX
 PH
 FT Key Location/Qualifiers
 FT Misc-difference 13..45
 FT /label= loop 1 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 7-23 of AAW14499 is deleted or
 FT replaced with a fragment corresponding to
 FT this fragment; claim 33"
 FT
 FT Misc-difference 14..40
 FT /label= loop 1 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 8-18
 FT of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim
 FT 42
 FT
 FT 35" Misc-difference 28..42
 FT /label= loop 1 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 12-19 of AAW14499 is deleted or
 FT replaced with a fragment corresponding to
 FT this fragment; claim 30"
 FT
 FT Misc-difference 32..38
 FT /label= loop 1 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 14-15
 FT of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim
 FT 34
 FT
 FT 32" Misc-difference 66..84
 FT /label= loop 2 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 44-57 of AAW14499 is deleted or
 FT replaced with a fragment corresponding to
 FT this fragment; claim 18"
 FT
 FT Misc-difference 70..78
 FT /label= loop 2 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 48-51
 FT of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim
 FT 20
 FT
 FT 20" Misc-difference 98..210
 FT /label= loop 3 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 117-185 of AAW14499 is deleted
 FT or replaced with a fragment corresponding to
 FT this fragment; claim 24"
 FT
 FT Misc-difference 102..206
 FT /label= loop 3 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 121-181
 FT of AAW14499 is deleted or replaced with a

FT fragment corresponding to this fragment; claim
 FT 26" Misc-difference 121..181
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to this fragment is deleted or
 FT replaced with a fragment corresponding to
 FT 102-206 of AAW14499; claim 41"
 FT
 FT Misc-difference 121..174
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to this
 FT fragment is deleted or replaced with a fragment
 FT corresponding to 102-199 of AAW14499; claim 42"
 FT
 FT Misc-difference 165..177
 FT /label= loop 3 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 195-202 of AAW14499 is deleted
 FT or replaced with a fragment corresponding to
 FT this fragment; claim 21"
 FT
 FT Misc-difference 166..173
 FT /label= loop 3 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 196-198
 FT of AAW14499 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim
 FT 23
 FT
 FT 23" Misc-difference 181..184
 FT /note= "an amino acid fragment corresponding to this
 FT region is deleted from the parent sequence of a
 FT variant Fungamyl; claim 43"
 FT
 FT Misc-difference 291..313
 FT /label= loop 8 modification region
 FT /note= "at least one amino acid residue of a parent
 FT alpha-amylase (used as a template for a variant)
 FT corresponding to 322-346 of AAW14498 is deleted
 FT or replaced with a fragment corresponding to
 FT this fragment; claim 36"
 FT
 FT Misc-difference 297..313
 FT /label= loop 8 modification region
 FT /note= "preferred region where at least one amino acid
 FT residue of a parent alpha-amylase (used as a
 FT template for a variant) corresponding to 325-345
 FT of AAW14498 is deleted or replaced with a
 FT fragment corresponding to this fragment; claim
 FT 38"
 FT
 FT WO9623874-A1.
 FT
 FT PD 08-AUG-1996.
 FT
 FT PF 05-FEB-1996; 96WO-DK00057.
 FT
 FT XX 10-NOV-1995; 95DK-0001256.
 FT PR 03-FEB-1995; 95DK-0000128.
 FT PR 23-OCT-1995; 95DK-0001192.
 FT
 FT (NOVO) NOVO-NORDISK AS.
 FT
 FT Bisgaard-frantzen H, Borchert TV, Svendsen A;
 FT WPI: 1996-371424/37.
 FT
 FT DX Alpha-amylase variants and methods of production - have altered
 FT PT properties such as calcium dependency, substrate binding and
 FT PT stability
 FT
 FT PS Disclosure: Page 87-88; 171pp; English.
 FT
 FT XX The present sequence is the mature Aspergillus oryzae alpha-amylase (A).
 FT CC Variants of parent termamyl- and fungamyl-like alpha-amylases (and
 FT CC methods of constructing them) are claimed. Examples of variants are
 FT CC featured above. The variants have altered properties such as calcium

XX XX

XX
XX

```

KW Alpha-amylase; thermostable enzyme; baking; Aspergillus niger.
XX
OS Thermomyces lanuginosus CBS 224.63.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= sig_peptide
XX
PN M09601323-A1.
XX
PD 18-JAN-1996.
XX
PE 03-JUL-1995; 95WO-EP02607.
XX
PR 04-JUL-1994; 94GB-0013419.
XX
PA (DANIT-) DANISCO AS.
XX
PL Michelsen B, Rasmussen P;
XX
DR WPI; 1996-087673/09.
DR N-PSDB; AAT10562.
XX
PT Thermophilic alpha-amylase with activity range of 60-80 degrees C -
PT derived from Thermomyces lanuginosus, useful in the prepn. of
PT foodstuffs and bakery prods. esp. bread
XX
PS Claim 3; Page 36-38; 94pp; English.
XX
CC A thermostable alpha-amylase (AAR88212) of Thermomyces lanuginosus
CC CBS 224.63 is expressed in Aspergillus niger hosts using a gene
CC (AAT10562) isolated from a T. lanuginosus gene library. The
CC recombinant enzyme (54-60 kDa) shows optimal activity at
CC 60-70 deg and pH 5.8-6, has a pI of 3.7 and is active at 60-80
CC deg.
XX
SQ Sequence 493 AA;

```

```

Query Match 88.5%; Score 69; DB 17; Length 493;
Best Local Similarity 76.9%; Pred. No. 0.0094;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AYHGWOODIYSL 13
:|||||:|||||
Db 98 sygywgkdlysl 110

```

```

RESULT 8
AAR79025
ID AAR79025 standard; protein; 478 AA.
XX
AC AAR79025;
XX
DT 22-MAR-1996 (first entry)
XX
DE Mature taka-amylase A.
XX
KW Wild type; neopullulanase; B. steatothermophilus; mutant;
KW food industry; modification; hydrophobicity; replacement
KW insertion; deletion.
XX
OS Aspergillus oryzae.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
PN JP07177891-A.
XX
PD 18-JUL-1995.

```

```

XX 31-OCT-1994; 94JP-0288658.
XX
PR 12-NOV-1993; 93JP-0306096.
XX
PA (EZAK ) EZAKI GLICO CO.
XX (NIDE ) NEC CORP.
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site - Increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
PS Disclosure; Page 10-11; 18pp; English.
XX
CC This sequence represents the mature form of taka-amylase from A. oryzae.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the
CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC hydrophobic amino acid from the selected site. The method was used
CC to produce neopullulanases Y377F, S422Y and M375L.
XX
SQ Sequence 478 AA;

```

```

Query Match 79.5%; Score 62; DB 16; Length 478;
Best Local Similarity 84.6%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 AYHGWOODIYSL 13
||| ||| ||| |||
Db 78 aytgywgtdlysl 90

```

```

RESULT 9
AAR24136
ID AAR24136 standard; Protein; 468 AA.
XX
AC AAR24136;
XX
DT 17-NOV-1992 (first entry)
XX
DE Alpha-amylase variant with leucine at position 84.
XX
KW Mutant; maltose; malto-oligosaccharides; Saccharomycopsis; fibuligera;
KW polymerisation; DP; transglycosylifer.
XX
OS Saccharomycopsis fibuligera.
XX
FH Key Location/Qualifiers
FT Misc-difference 84 /note="mutation to leucine"
FT
XX
PN JP04108386-A.
XX
PD 09-APR-1992.
XX
PE 28-AUG-1990; 90JP-0226112.
XX
PR 28-AUG-1990; 90JP-0226112.
XX
PA (AGEN ) AGENCY OF IND SCI &.
XX
DR WPI; 1992-171652/21.
DR N-PSDB; AAO25079.
XX
PT Variant alpha-amylase gene for mfr. of malto-oligosaccharide(s) -
PT is obtd. by mutating the nucleotide at position 329 of the

```

PT Saccharomycopsis fibuligera wild-type sequence to thymine.
 XX
 PS Claim 2; Fig 1; 10pp; Japanese.
 XX

CC The variant alpha amylase was obtd. from a gene contg. a mutation
 CC of the 329th nucleotide of the alpha amylase gene of Saccharomycopsis
 CC fibuligera to T. This mutation results in substitution of the wild-
 CC type amino acid at position 84 of alpha-amylase by leucine. The
 CC variant alpha-amylase is high in transglycosylating activity.
 CC variant alpha-amylase may be used to prepare malto-oligosaccharides
 CC with a degree of polymerisation (DP) of at least 7, by inversion of
 CC the malto-oligosaccharide.
 XX

SQ Sequence 468 AA;

Query Match 74.4%; Score 58; DB 13; Length 468;
 Best Local Similarity 61.5%; Pred. No. 0.45;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGYYQODIYSL 13
 |||||:|:|:
 Db 79 ayhgywmkniyki 91

RESULT 10

ID AAP70571 standard; Protein; 494 AA.

AC AAP70571;

DT 03-MAY-1991 (first entry)

DE Alpha-amylase gene product from plasmid pslf alpha 1.

XX Amylase.

OS Saccharomycetes fibuligera HUT7212.

PN JP62104576-A.

PD 15-MAY-1987.

PF 31-OCT-1985; 85JP-0244892.

PR 31-OCT-1985; 85JP-0244892.

PA (FUKU/) FUKUI S.

DR WPI; 1987-173694/25.

DR N-PSDB; AAN70916.

PT Amylase prodn. - comprises culturing microorganism transformed
 PT with vector deoxyribonucleic acid, accumulating and collecting
 PT amylase

PS Disclosure; Fig 1; 14pp; Japanese.

XX

CC The product is an amylase biosynthetic enzyme encoded by a plasmid
 CC which may be used to transform an E.coli expression system for

CC the stable production of amylase, useful in ethanol fermentation.

XX

SQ Sequence 494 AA;

Query Match 74.4%; Score 58; DB 8; Length 494;
 Best Local Similarity 61.5%; Pred. No. 0.48;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGYYQODIYSL 13
 |||||:|:|:
 Db 105 ayhgywmkniyki 117

RESULT 11

ID AAR63184 standard; Protein; 468 AA.

AC AAR63184;

DT 16-JUN-1995 (first entry)

DE Variant alpha amylase deriv. from Saccharomycopsis fibuligera.

XX alpha amylase; carbohydrate hydrolase; increased activity;

KM tyrosine residue; enzyme centre; mass production; oligosaccharide;

XX variant; cyclomalto-dextrin glucanotransferase.

OS Saccharomycopsis fibuligera.

PN JP06253836-A.

PD 13-SEP-1994.

PF 04-MAR-1993; 93JP-0069303.

PR 04-MAR-1993; 93JP-0069303.

PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.

DR WPI; 1994-328987/41.

DR N-PSDB; AA077665.

PT Variant carbohydrate hydrolase(s) with increased activity -

PT consists of e.g. alpha-amylase with tyrosine residue in enzyme

PT centre, useful for mass-prodn. of oligosaccharide(s)

XX

PS Example 1; Page 13-15; 27pp; Japanese.

CC AAR63184-7 show the amino acid sequences of variant alpha amylases,

CC composed by substituting bases 247-249 of the structural gene (see

CC AA077665-8) region, with TTC, TGG, CTC or AAC. These substitutions

CC result in the 83rd amino acid residue (tyrosine) of the wild type

CC sequence being changed to phenylalanine, tryptophan, leucine or

CC asparagine respectively. The substituted amino acid is present in

CC the active site of the enzyme and confers increased activity on the

CC enzyme. The variants are useful for the mass production of a

CC oligosaccharides. (see AAR63188 for the variant structure of a

CC cyclomalto-dextrin glucanotransferase).

XX

SQ Sequence 468 AA;

Query Match 69.2%; Score 54; DB 15; Length 468;
 Best Local Similarity 53.8%; Pred. No. 1.9;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGYYQODIYSL 13
 |||||:|:|:
 Db 79 ayhgywmkniyki 91

RESULT 12

ID AAR63185 standard; Protein; 468 AA.


```

AC  AAR63185;
XX
XX  16-JUN-1995 (first entry)
DE  Variant alpha amylase deriv. from Saccharomycopsis fibuligera.
XX
XX  alpha amylase; carbohydrate hydrolase; increased activity;
XX  tyrosine residue; enzyme centre; mass production; oligosaccharide;
XX  variant; cyclomaltoedextrin glucanotransferase.
OS  Saccharomycopsis fibuligera.
FH  Key
FT  Misc-difference 83
FT  /label= Y83W
XX
XX  JP06253836-A.
XX
XX  13-SEP-1994.
XX
XX  04-MAR-1993; 93JP-0069303.
XX
XX  04-MAR-1993; 93JP-0069303.
XX
XX  (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
XX  WPI: 1994-328987/41.
XX  N-PSDB: AAQ77666.
XX
XX  Variant carbohydrate hydrolase(s) with increased activity -
XX  consists of e.g. alpha-amylase with tyrosine residue in enzyme
XX  centre, useful for mass-prodn. of oligosaccharide(s)
XX
XX  Example 1; Page 15-17; 27pp; Japanese.
XX
XX  AAR63184-7 show the amino acid sequences of variant alpha amylases,
XX  composed by substituting bases 247-249 of the structural gene (see
XX  AAQ77665-8) region, with TTC, TGC, CTC or AAC. These substitutions
XX  result in the 83rd amino acid residue (tyrosine) of the wild type
XX  sequence being changed to phenylalanine, tryptophan, leucine or
XX  asparagine respectively. The substituted amino acid is present in
XX  the active site of the enzyme and confers increased activity on the
XX  enzyme. The variants are useful for the mass production of
XX  oligosaccharides. (see AAR63188 for the variant structure of a
XX  cyclomaltoedextrin glucanotransferase).
XX
XX  Sequence 468 AA;
XX
XX  Query Match 67.9%; Score 53; DB 15; Length 468;
XX  Best Local Similarity 53.8%; Pred. No. 2.7;
XX  Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX
OY  1 AYHGYYODDIYSL 13
XX  |||||:||||:
DB  79 ayhgywmknlyki 91

RESULT 13
AAR63186
ID  AAR63186 standard; Protein; 452 AA.
XX
XX  AAR63186;
XX
XX  16-JUN-1995 (first entry)
XX
XX  Variant alpha amylase deriv. from Saccharomycopsis fibuligera.
XX
XX  alpha amylase; carbohydrate hydrolase; increased activity;
XX  tyrosine residue; enzyme centre; mass production; oligosaccharide;
XX  variant; cyclomaltoedextrin glucanotransferase.
OS  Saccharomycopsis fibuligera.

```

```

XX
XX  Key Location/Qualifiers
FH  Misc-difference 83
FT  /label= Y83L
XX
XX  JP06253836-A.
XX
XX  13-SEP-1994.
XX
XX  04-MAR-1993; 93JP-0069303.
XX
XX  04-MAR-1993; 93JP-0069303.
XX
XX  (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
XX  WPI: 1994-328987/41.
XX  N-PSDB: AAQ77667.
XX
XX  Variant carbohydrate hydrolase(s) with increased activity -
XX  consists of e.g. alpha-amylase with tyrosine residue in enzyme
XX  centre, useful for mass-prodn. of oligosaccharide(s)
XX
XX  Example 1; Page 18-20; 27pp; Japanese.
XX
XX  AAR63184-7 show the amino acid sequences of variant alpha amylases,
XX  composed by substituting bases 247-249 of the structural gene (see
XX  AAQ77665-8) region, with TTC, TGC, CTC or AAC. These substitutions
XX  result in the 83rd amino acid residue (tyrosine) of the wild type
XX  sequence being changed to phenylalanine, tryptophan, leucine or
XX  asparagine respectively. The substituted amino acid is present in
XX  the active site of the enzyme and confers increased activity on the
XX  enzyme. The variants are useful for the mass production of
XX  oligosaccharides. (see AAR63188 for the variant structure of a
XX  cyclomaltoedextrin glucanotransferase).
XX
XX  Sequence 452 AA;
XX
XX  Query Match 64.1%; Score 50; DB 15; Length 452;
XX  Best Local Similarity 53.8%; Pred. No. 7.6;
XX  Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
XX
OY  1 AYHGYYODDIYSL 13
XX  |||||:||||:
DB  79 ayhgywmknlyki 91

RESULT 14
AAR63186
ID  AAR63186 standard; Protein; 511 AA.
XX
XX  AAR63186;
XX
XX  12-OCT-1990 (first entry)
XX
XX  Recombinant alpha-amylase.
XX
XX  Alpha amylase; yeast; brewing; baking.
XX
XX  Schwanmionycos castelli ATCC 26076.
XX
XX  EP257115-A.
XX
XX  02-MAR-1988.
XX
XX  21-AUG-1986; 86EP-0111586.
XX
XX  21-AUG-1986; 86EP-0111586.
XX
XX  (HETB ) HEINCKEN TECHNISCH BEHEER BV.
XX
XX  Strasser A, Martens FB, Dohmen J, Hollenberg CP;
XX

```

DR WPI: 1988-057532/09.
 DR N-PSDB; AAN81477.
 XX
 XX Prodn. of amylolytic enzymes by transformed microorganisms -
 PT comprises use of recombinant DNA technology to give transformant
 PT having DNA from donor yeast.
 XX
 PS Disclosure: ; p; English.
 XX
 CC The sequence encoding the protein can be used to construct an
 CC expression vector for the prodn. of alpha amylase in yeast cells.
 CC The transformants and recombinant enzyme produced can be used in
 CC fermentation processes e.g in baking and brewing. The transformed
 CC yeasts are esp. suitable for prodn. of low carbohydrate beers.
 XX
 SQ Sequence 511 AA;

Query Match 64.1%; Score 50; DB 9; Length 511;
 Best Local Similarity 61.5%; Pred. No. 8.7;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 OY 1 AYHGWOODIYSL 13
 |||||::|
 Db 114 ayhgywmknidel 126

RESULT 15
 AAP81180
 ID AAP81180 standard; protein; 511 AA.
 XX
 AC AAP81180;
 XX
 DT 12-DEC-1990 (first entry)
 XX
 DE Sequence of alpha-amylase encoded by genomic DNA of Schwanniomycetes
 DE castellii.
 XX
 KW Brewing; beer; breadmaking; biomass.
 XX
 OS Schwanniomycetes castellii.
 XX
 PN EP260404-A.
 PD 23-MAR-1988.
 XX
 PF 17-JUL-1987; 87EP-0110370.
 XX
 PR 17-JUL-1987; 87EP-0110370.
 XX
 PA (HETB) HEINEKEN TEC BEHEER NV.
 XX
 PI Strasser A, Martens FB, Dohmen J, Hollenberg CP;
 DR WPI: 1988-078794/12.
 DR N-PSDB; AAN81525.
 XX
 PT Polypeptide prodn. from cells transformed with yeast DNA -
 PT esp. coding alpha amylase or glucoamylase, able to convert
 PT starch to ethanol
 XX
 PS Claim 18; Fig 2A-2C; 74pp; English.
 XX
 CC A cosmid library is constructed from genomic DNA of ATCC26076 and this
 CC screened for inserts contg. the AA gene by ability to transform
 CC S.cerevisiae GRF 18 to histidine prototrophy. One positive cosmid
 CC (32kb; pYcl) was digested with EcoRI, religated and used to transform
 CC E.coli JA221. Positive transformants contain pYcl-alpha, which includes
 CC a 5kb EcoRI fragment contg. the gene in AAN81525.
 XX
 SQ Sequence 511 AA;

Query Match 64.1%; Score 50; DB 9; Length 511;
 Best Local Similarity 61.5%; Pred. No. 8.7;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 OY 1 AYHGWOODIYSL 13
 |||||::|
 Db 114 ayhgywmknidel 126

Search completed: June 13, 2002, 08:53:25
 Job time: 275 sec

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:31 ; Search time 57.95 Seconds
(without alignments)
3.793 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Perfect score: 44

Sequence: 1 NAKKSTSGS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgcn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgcn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgcn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgcn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgcn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgcn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	478	1	US-08-720-899-7
2	44	100.0	478	1	US-08-459-610-7
3	44	100.0	478	2	US-08-343-804-7
4	44	100.0	478	2	US-08-339-715A-2
5	44	100.0	478	2	US-08-600-908A-10
6	44	100.0	478	3	US-08-683-838A-10
7	44	100.0	478	4	US-09-182-859-7
8	31	70.5	367	3	US-09-226-741-3
9	31	70.5	367	3	US-09-226-741-1
10	30	68.2	606	4	US-09-041-236-4
11	29	65.9	161	1	US-08-145-995A-13
12	29	65.9	161	1	US-08-451-747-13
13	29	65.9	161	3	US-09-134-852-13
14	29	65.9	870	2	US-09-010-928B-2
15	29	65.9	2089	1	US-08-418-893D-23
16	29	65.9	2089	1	US-08-418-893D-24
17	28	63.6	126	2	US-08-482-728A-10
18	28	63.6	127	4	US-09-160-246-16
19	28	63.6	208	1	US-08-142-897-7
20	28	63.6	248	3	US-08-750-145A-20
21	28	63.6	248	3	US-08-975-698A-24
22	28	63.6	248	4	US-09-417-090-24
23	28	63.6	248	4	US-09-727-578-24
24	28	63.6	280	4	US-09-160-246-14
25	28	63.6	314	1	US-07-704-288C-7
26	28	63.6	314	1	US-08-379-259-7
27	28	63.6	377	1	US-08-525-697-2

28	28	63.6	466	4	US-08-868-435-2	Sequence 2, Appli
29	28	63.6	466	4	US-08-744-231-2	Sequence 2, Appli
30	28	63.6	597	1	US-08-844-274-12	Sequence 12, Appli
31	28	63.6	1188	4	US-08-201-697-4	Sequence 4, Appli
32	28	63.6	1437	3	US-09-061-400-2	Sequence 2, Appli
33	28	63.6	1437	3	US-09-061-273-2	Sequence 2, Appli
34	28	63.6	1453	4	US-08-843-459A-2	Sequence 2, Appli
35	28	63.6	1912	4	US-08-913-832A-2	Sequence 2, Appli
36	27	61.4	152	2	US-08-722-050-12	Sequence 12, Appli
37	27	61.4	154	4	US-08-679-493A-211	Sequence 12, Appli
38	27	61.4	185	2	US-08-691-814B-12	Sequence 12, Appli
39	27	61.4	245	1	US-08-317-707-2	Sequence 2, Appli
40	27	61.4	245	2	US-08-514-921-2	Sequence 2, Appli
41	27	61.4	245	2	US-08-514-921-4	Sequence 4, Appli
42	27	61.4	245	2	US-08-514-921-6	Sequence 6, Appli
43	27	61.4	245	2	US-08-514-921-8	Sequence 8, Appli
44	27	61.4	353	1	US-08-073-807A-17	Sequence 17, Appli
45	27	61.4	380	1	US-08-073-807A-18	Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7

Query Match 100.0%; Score 44; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 |||||
 Db 260 NAFKSTSGS 268

RESULT 2
 US-08-459-610-7
 ; Sequence 7, Application US/08459610
 ; Patent No. 5801043
 ; GENERAL INFORMATION:
 ; APPLICANT: Bisgaard-Frantzen, Henrik
 ; APPLICANT: Borchert, Torben Vedel
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Thellersen, Marianne
 ; TITLE OF INVENTION: AMYLASE VARIANTS
 ; NUMBER OF SEQUENCES: 38
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/459,610
 ; FILING DATE: 02-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/343,804
 ; FILING DATE: 22-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney Dr., Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4054.214-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 478 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-459-610-7

Query Match 100.0%; Score 44; DB 1; Length 478;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 |||||
 Db 260 NAFKSTSGS 268

RESULT 3
 US-08-343-804-7
 ; Sequence 7, Application US/08343804
 ; Patent No. 5830837
 ; GENERAL INFORMATION:
 ; APPLICANT: Bisgaard-Frantzen, Henrik
 ; APPLICANT: Borchert, Torben Vedel
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Thellersen, Marianne
 ; APPLICANT: Van der Zee, Pia
 ; TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 58308370 No. 5830837disk of No. 5830837th America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/343,804
 ; FILING DATE: 22-NOV-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney Dr., Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4054.214-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 478 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-343-804-7

Query Match 100.0%; Score 44; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 |||||
 Db 260 NAFKSTSGS 268

RESULT 4
 US-08-339-715A-2
 ; Sequence 2, Application US/08339715A
 ; Patent No. 5965442
 ; GENERAL INFORMATION:
 ; APPLICANT: kaneko, Hiroki
 ; APPLICANT: Takada, Toshikazu
 ; APPLICANT: Shimada, Jiro
 ; APPLICANT: Kuriki, Takashi
 ; APPLICANT: Yanase, Michiyo
 ; APPLICANT: Takata, Hiroki
 ; APPLICANT: Okada, Shigetaka
 ; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: POPHAM, HARK, SCHNOBRICH & KAUFMAN, LTD.
 ; STREET: Metropolitan Square Building, Suite 800, 1450
 ; STREET: G. Street
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/339,715A
 ; FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 44; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
Db 260 NAFKSTSGS 268

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-600-908A-10

Query Match 100.0%; Score 44; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
Db 260 NAFKSTSGS 268

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-683-838A-10

Query Match 100.0%; Score 44; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
Db 260 NAFKSTSGS 268

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgaard-Frantzen, Henrik

FILE OF INVENTION: Alpha-Amylase Mutants
FILE REFERENCE: 4796.204-US
CURRENT APPLICATION NUMBER: US/09/182,859
CURRENT FILING DATE: 1998-10-29
EARLIER APPLICATION NUMBER: 0515/96
EARLIER FILING DATE: 1996-04-30
EARLIER APPLICATION NUMBER: 0712/96
EARLIER FILING DATE: 1996-06-28
EARLIER APPLICATION NUMBER: 0775/96
EARLIER FILING DATE: 1996-07-11
EARLIER APPLICATION NUMBER: 1263/96
EARLIER FILING DATE: 1996-11-08
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 478
TYPE: PRT
ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 44; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||||
DB 260 NAFKSTSGS 268

RESULT 8
US-09-226-741-3
Sequence 3, Application US/09226741
Patent No. 6100037
GENERAL INFORMATION:
APPLICANT: Phillips, Stephen C.
APPLICANT: Lanfear, Jerry
APPLICANT: Fawcett, Lindsay
APPLICANT: Bandman, Olga
TITLE OF INVENTION: HUMAN CYCLIC NUCLEOTIDE PDES
FILE REFERENCE: PF-0623 US
CURRENT APPLICATION NUMBER: US/09/226,741
CURRENT FILING DATE: 1999-01-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 367
TYPE: PRT
ORGANISM: HOMO SAPIENS
FEATURE:
OTHER INFORMATION: HSPDE10A2
US-09-226-741-3

Query Match 70.5%; Score 31; DB 3; Length 367;
Best Local Similarity 66.7%; Pred. No. 83;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||||
DB 269 NAFKSTSGS 277

RESULT 9
US-09-226-741-1
Sequence 1, Application US/09226741
Patent No. 6100037
GENERAL INFORMATION:
APPLICANT: Phillips, Stephen C.
APPLICANT: Lanfear, Jerry
APPLICANT: Fawcett, Lindsay
APPLICANT: Bandman, Olga
TITLE OF INVENTION: HUMAN CYCLIC NUCLEOTIDE PDES

FILE OF INVENTION: PF-0623 US
CURRENT APPLICATION NUMBER: US/09/226,741
CURRENT FILING DATE: 1999-01-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 490
TYPE: PRT
ORGANISM: HOMO SAPIENS
FEATURE:
OTHER INFORMATION: HSPDE10A1
US-09-226-741-1

Query Match 70.5%; Score 31; DB 3; Length 490;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||||
DB 269 NAFKSTSGS 277

RESULT 10
US-09-041-236-4
Sequence 4, Application US/09041236
Patent No. 6225285
GENERAL INFORMATION:
APPLICANT: Luo, Yaling
APPLICANT: Xiomel, Xu
TITLE OF INVENTION: Semaphorin K1
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,236
FILING DATE: March 11, 1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: EXEL98-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-236-4

Query Match 68.2%; Score 30; DB 4; Length 606;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||||
DB 167 NAFKSTSGS 175


```
RESULT 11
US-08-145-995A-13
; Sequence 13, Application US/08145995A
; Patent No. 5482850
; GENERAL INFORMATION:
; APPLICANT: CARLOW, CLOTTILDE K.S.
; APPLICANT: PAGE, ANTONY
; TITLE OF INVENTION: METHOD FOR IDENTIFYING ANTI-PARASITIC
; TITLE OF INVENTION: COMPOUNDS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,995A
; FILING DATE: 29-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 43406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-145-995A-13

Query Match 65.9%; Score 29; DB 1; Length 161;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
Db 99 NAGKNTNGS 107

RESULT 12
US-08-451-747-13
; Sequence 13, Application US/08451747
; Patent No. 5821107
; GENERAL INFORMATION:
; APPLICANT: CARLOW, CLOTTILDE K.S.
; APPLICANT: PAGE, ANTONY
; TITLE OF INVENTION: METHOD FOR IDENTIFYING ANTI-PARASITIC
; TITLE OF INVENTION: COMPOUNDS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,747
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/145,995
; FILING DATE: 29-OCT-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-046-DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-451-747-13

Query Match 65.9%; Score 29; DB 2; Length 161;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
Db 99 NAGKNTNGS 107

RESULT 13
US-09-134-852-13
; Sequence 13, Application US/09134852
; Patent No. 6127148
; GENERAL INFORMATION:
; APPLICANT: CARLOW, CLOTTILDE K.S.
; APPLICANT: PAGE, ANTONY
; TITLE OF INVENTION: METHOD FOR IDENTIFYING ANTI-PARASITIC
; TITLE OF INVENTION: COMPOUNDS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,852
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/145,995
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 43406
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-134-852-13

Query Match 65.9%; Score 29; DB 3; Length 161;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||:|
DB 99 NAGKNTGS 107

RESULT 14
US-09-010-928B-2
; Sequence 2, Application US/09010928B
; Patent No. 5994099
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V
; APPLICANT: Hayashi, Cheryl Y
; TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
; TITLE OF INVENTION: CODING THEREFOR
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOIASCH & BIRCH
; STREET: 8110 GATEHOUSE RD. SUITE 500E
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,928B
; FILING DATE: 22-JAN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28977
; REFERENCE/DOCKET NUMBER: 1447-109P
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 870 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-010-928B-2

Query Match 65.9%; Score 29; DB 2; Length 870;
Best Local Similarity 66.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||:|
DB 49 NALMSTGS 57

RESULT 15
US-08-418-893D-23
; Sequence 23, Application US/08418893D

; Patent No. 5559220
; GENERAL INFORMATION:
; APPLICANT: ROESSLER, PAUL G
; APPLICANT: OHLROGGE, JOHN B
; TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A
; TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTETRA CRYPTIC
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY
; STREET: 1617 Cole Blvd.
; CITY: Golden
; STATE: CO
; COUNTRY: USA
; ZIP: 80401-3393
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,893D
; FILING DATE: April 7, 1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,938
; FILING DATE: September 14, 1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: O'CONNOR, EDNA
; REGISTRATION NUMBER: 29,252
; REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303-231-1000
; TELEFAX: 303-231-1098
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2089 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-418-893D-23

Query Match 65.9%; Score 29; DB 1; Length 2089;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 FKSTSGS 9
|||:|
DB 502 FKPTSGS 508

Search completed: June 13, 2002, 08:54:32
Job time: 317 sec

Fri Jun 14 10:36:46 2002

us-09-710-339-2_copy_280_288.ra1

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:12 ; Search time 125.81 Seconds
(without alignments)
12.375 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Perfect score: 44

Sequence: 1 MAFKSTSGS 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phase:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	100.0	499	3	Q96TH4	Q96TH4 aspergillus
2	40	90.9	623	3	Q9UV09	Q9UV09 emericella
3	37	84.1	2276	2	Q93TV6	Q93TV6 staphylococ
4	34	77.3	102	3	P87271	P87271 saccharomyc
5	34	77.3	902	6	Q9BDY9	Q9BDY9 echinops te
6	34	77.3	915	6	Q9BDR8	Q9BDR8 tenrec ecan
7	34	77.3	1146	6	Q46488	Q46488 alouatta se
8	33	75.0	297	16	Q9KBL4	Q9KBL4 bacillus ha
9	33	75.0	490	3	Q9UV07	Q9UV07 emericella
10	33	75.0	631	3	Q92394	Q92394 cryptococcu
11	33	75.0	688	5	Q917J5	Q917J5 drosophila
12	33	75.0	690	5	Q9NH72	Q9NH72 drosophila
13	33	75.0	690	5	Q9VFT4	Q9VFT4 drosophila
14	32	72.7	100	17	Q9VAT6	Q9VAT6 aeropyrum p
15	32	72.7	196	11	Q9BOM8	Q9BOM8 mus musculu
16	32	72.7	220	15	Q900X2	Q900X2 human immun

17	32	72.7	399	13	Q98SE8	Q98SE8 gallus gall
18	32	72.7	427	17	Q97C83	Q97C83 thermoplasma
19	32	72.7	513	10	Q81992	Q81992 hordium vul
20	32	72.7	513	10	Q40029	Q40029 hordium vul
21	32	72.7	681	5	Q20353	Q20353 caenorhabdi
22	32	72.7	760	3	Q9P3E0	Q9P3E0 neurospora
23	32	72.7	937	4	Q9H3Q4	Q9H3Q4 homo sapien
24	32	72.7	1198	5	Q9U5Y2	Q9U5Y2 drosophila
25	32	72.7	1200	5	Q9V8U8	Q9V8U8 drosophila
26	32	72.7	1200	5	Q9U531	Q9U531 drosophila
27	32	72.7	1200	5	Q9U472	Q9U472 drosophila
28	32	72.7	1638	4	Q9Y6U4	Q9Y6U4 homo sapien
29	32	72.7	1649	16	Q9CF22	Q9CF22 lactococcus
30	32	72.7	1712	4	Q9Y6U5	Q9Y6U5 homo sapien
31	31	70.5	194	10	Q93ZG2	Q93ZG2 arabidopsis
32	31	70.5	221	8	Q96G11	Q96G11 gomphus flo
33	31	70.5	234	16	Q92E22	Q92E22 listeria in
34	31	70.5	419	13	Q92043	Q92043 croctalus at
35	31	70.5	471	10	Q9XIB2	Q9XIB2 arabidopsis
36	31	70.5	480	4	Q9H0D4	Q9H0D4 homo sapien
37	31	70.5	490	4	Q9NY45	Q9NY45 homo sapien
38	31	70.5	492	4	Q9NSN4	Q9NSN4 homo sapien
39	31	70.5	492	16	Q9CE95	Q9CE95 lactococcus
40	31	70.5	562	11	Q9WVG4	Q9WVG4 rattus norv
41	31	70.5	572	5	Q26042	Q26042 penaeus jap
42	31	70.5	576	4	Q9HB46	Q9HB46 homo sapien
43	31	70.5	595	10	Q9AR86	Q9AR86 populus x c
44	31	70.5	607	5	Q9U5H3	Q9U5H3 caenorhabdi
45	31	70.5	621	11	Q61866	Q61866 mus musculu

ALIGNMENTS

RESULT 1
ID Q96TH4 PRELIMINARY; PRT; 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
GN AMYA OR AMY1.
OS Aspergillus oryzae, and
OS Aspergillus oryzae, and
OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062, 5059;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A.oryzae; STRAIN=RIB40;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
RA Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator
RT gene, amyR, involved in the amylolytic gene expression in Aspergillus
RT oryzae.";
RT Biosci. Biotechnol. Biochem. 64:816-827(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=A.flavus; STRAIN=86-10D;
RA Fakhoury A.M., Woloshuk C.P.;
RT "Amy1, the alpha-amylase gene of Aspergillus flavus: Involvement in
RT aflatoxin biosynthesis in maize kernels.";
RT Phytopathology 89:908-914(1999).
DR EMBL; AB021876; BAA95703.1; -
DR EMBL; AF139925; AAF14264.1; -
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match 100.0%; Score 44; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 |||||
 DB 281 NAFKSTSGS 289

RESULT 2
 ID Q9UV09 PRELIMINARY; PRT; 623 AA.
 AC Q9UV09;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ALPHA-AMYLAASE.
 GN AMYB.
 OS *Emmericella nidulans* (*Aspergillus nidulans*).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutiales; Trichocomaceae; *Emmericella*.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boase N.A., Murphy R.L., Kelly J.M.;
 RT "Characterisation and regulation of an alpha-amyase gene in
Aspergillus nidulans.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR HSP; p56271; 2AAA
 DR EMBL; AF208224; AAF17100.1; -
 DR InterPro; IPR000461; Alpha_amyase.
 DR InterPro; IPR002044; CBD_4.
 DR Pfam; PF00128; alpha-amyase; 1.
 DR Pfam; PF00686; CBD_4; 1.
 DR PRINTS; PR00110; ALPHAMYLASE.
 DR Prodom; PD001568; CBD_4; 1.
 SO SEQUENCE 623 AA; 68598 MW; 822F002C37F5A9A4 CRC64;

Query Match 90.9%; Score 40; DB 3; Length 623;
 Best Local Similarity 88.9%; Pred. No. 4.1;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NAFKSTSGS 9
 |||||
 DB 279 NAFKSTSGS 287

RESULT 3
 ID Q93TY6 PRELIMINARY; PRT; 2276 AA.
 AC Q93TY6;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE BIOFILM-ASSOCIATED SURFACE PROTEIN.
 GN BAP.
 OS *Staphylococcus aureus*.
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;
 OC *Bacillus*/Staphylococcus group; *Staphylococcus*.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V329;
 RX MEDLINE=21189316; PubMed=11292810;
 RA Cucarella C., Solano C., Valle J., Amorena B., Lasa I.,
 RA Penades J.R.;
 RT "Bap, a *Staphylococcus aureus* Surface Protein Involved in Biofilm
 Formation.";
 RL J. Bacteriol. 183:2888-2896(2001).
 DR EMBL; AF288402; AAK38834.1; -
 SO SEQUENCE 2276 AA; 238792 MW; 8623C60472CBOC0F CRC64;

Query Match 84.1%; Score 37; DB 2; Length 2276;
 Best Local Similarity 77.8%; Pred. No. 65;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NAFKSTSGS 9
 |||||
 DB 548 NAFKSTSGS 556

RESULT 4
 ID P87271 PRELIMINARY; PRT; 102 AA.
 AC P87271;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE YDR455CP.
 GN YDR455C.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; *Saccharomyces*.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dietrich F.S.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 DR [2]
 RP SEQUENCE FROM N.A.
 RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
 RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
 RA Hunik-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
 RA Moseedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Stroff N.,
 RA Winant A., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Jia Y., Cherry J.M.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U33007; AAB6488.1; -
 DR SGD; S0002863; YDR455C.
 SO SEQUENCE 102 AA; 10691 MW; 0EFC38552A82BB9 CRC64;

Query Match 77.3%; Score 34; DB 3; Length 102;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 FKSTSGS 9
 |||||
 DB 15 FKSTSGS 21

RESULT 5
 ID Q9BDT9 PRELIMINARY; PRT; 902 AA.
 AC Q9BDT9;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE BROCI (FRAGMENT)
 OS *Echinops telfairi* (Lesser hedgehog tenrec).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Insectivora; Tenrecidae; Tenrecinae; *Echinops*.
 OX NCBI_TaxID=9371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082081; PubMed=11214318;
 RA Maden O., Scally M., Donady C., Kao D., DeBry R., Adkins R.,
 RA Amrine H., Stanhope M., de Jong W., Springer M.;
 RT "Parallel adaptive radiations in two major clades of placental
 mammals.";
 RL Nature 409:610-614(2001).
 DR EMBL; AF284025; AAK15612.1; -
 DR InterPro; IPR002378; Brst_cancerI.

DR InterPro: IPR001993; Mitoch_carrier.
DR PRINTS: PR00493; BRSTCANCER1.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 902
SQ SEQUENCE 902 AA; 98813 MW; CD8E27EA1A3392CD CRC64;

Query Match
Best Local Similarity 77.3%; Score 34; DB 6; Length 902;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
Db 738 NAFKGTSSS 746

RESULT 6
Q9BDT8 PRELIMINARY; PRT; 915 AA.
AC Q9BDT8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BRCA1 (FRAGMENT).
OS Tenrec ecaudatus (tailless tenrec).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Tenrecidae; Tenrecinae; Tenrec.
OX NCBI_TaxID=94439;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21082081; PubMed=11214318;
RA Madden O., Scally M., Donady C., Kao D., Debry R., Adkins R.,
RA Amrine H., Stanhope M., de Jong W., Springer M.;
RT "Parallel adaptive radiations in two major clades of placental
mammals."
RL Nature 409:610-614(2001).
DR EMBL: AF284026; AAK15613.1; -
DR InterPro: IPR002378; Brstl_cancert.
DR PRINTS: PR00493; BRSTCANCER1.
FT NON_TER 1
FT NON_TER 915
SQ SEQUENCE 915 AA; 100159 MW; A9D62AA3E83E8F36 CRC64;

Query Match
Best Local Similarity 77.3%; Score 34; DB 6; Length 915;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
Db 743 NAFKGTSSS 751

RESULT 7
Q46488 PRELIMINARY; PRT; 1146 AA.
AC Q46488;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE BREAST CANCER TYPE 1 SUSCEPTIBILITY PROTEIN (BREAST AND OVARIAN CANCER
DE SUSCEPTIBILITY PROTEIN) (FRAGMENT).
GN BRCA1.
OS Alothatta seniculus (Howler monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae;
OC Alouatta.
OX NCBI_TaxID=9503;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98122577; PubMed=9462745;
RA Hacia J.G., Makalowski W., Edgemon K., Erdos M.R., Robbins C.M.,

RA Fodor S.P.A., Brody L.C., Collins F.S.;
RT "Evolutionary sequence comparisons using high-density oligonucleotide
RT arrays."
RL Nat. Genet. 18:155-158(1998).
CC -I- FUNCTION: NOT KNOWN, MAY REGULATE GENE EXPRESSION. MAY PLAY A ROLE
CC IN CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
DR EMBL: AF019079; AAC39587.1; -
DR InterPro: IPR002378; Brstl_cancert.
DR PRINTS: PR00493; BRSTCANCER1.
KM DNA-binding; Nuclear protein; Anti-oncogene.
FT NON_TER 1
FT DOMAIN 278 285 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 200 203 POLY-SER.
FT DOMAIN 815 818 POLY-SER.
FT DOMAIN 1001 1004 POLY-GLU.
FT DOMAIN 996 1005 LEUCINE-ZIPPER (POTENTIAL).
FT NON_TER 1146 1146
SQ SEQUENCE 1146 AA; 128119 MW; BFD07EBBF851890B CRC64;

Query Match
Best Local Similarity 77.3%; Score 34; DB 6; Length 1146;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
Db 809 NAFKGTSSS 817

RESULT 8
Q9KBL4 PRELIMINARY; PRT; 297 AA.
AC Q9KBL4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE TRANSMEMBRANE LIPOPROTEIN.
GN LpIC OR Bhl912.
GN Bacillus halodurans.
OS Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Holikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AP001513; BAB05631.1; -
DR InterPro: IPR000515; BPD_transp.
DR Pfam: PF00528; BPD_transp.1.
DR PROSITE: PS00402; BPD_TRANSF_LIN_MEMBER; UNKNOWN_1.
DR PROSITE: PS00402; BPD_TRANSF_LIN_MEMBER; UNKNOWN_1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 297 AA; 33588 MW; BC744C0A5B8FAEE CRC64;

Query Match
Best Local Similarity 75.0%; Score 33; DB 16; Length 297;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NAFKSTSG 8
Db 244 NAFKGTSG 251

RESULT 9
Q9UV07 PRELIMINARY; PRT; 490 AA.

```
AC Q9UV07;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE AMYA.
GN AMYA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiaceae; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RA SEQUENCE FROM N.A.
RA Rose N.A., Murphy R.M., Kelly J.M.;
RT "An amylase cluster in Aspergillus nidulans."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208225; AAF17103.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR000461; Alpha-amylase.
DR Pfam; PF00128; alpha-amylase_1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR PRINTS; PR00110; ALPHAMYLASE.
SQ SEQUENCE 490 AA; 54249 MW; A891C4ACEAB5305 CRC64;

Query Match
Best Local Similarity 75.0%; Score 33; DB 3; Length 490;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AEFSTSGS 9
DB 274 AEFSTSGS 281

RESULT 10
ID Q92394 PRELIMINARY; PRT; 631 AA.
AC Q92394;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR.
OS AMY-CS2.
OS Cryptococcus sp. S-2.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
OC mitospotic Hymenomycetes; Cryptococcus.
OX NCBI_TaxID=87049;
RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN-S-2;
RX MEDLINE=96433120; PubMed=8836148;
RA Iefuji H., Chino M., Kato M., Iimura Y.;
RT "Raw-starch-digesting and thermostable alpha-amylase from the yeast
RT Cryptococcus sp. S-2: purification, characterization, cloning and
RT sequencing."
RL Biochem. J. 318:989-996(1996).
DR EMBL; D83541; BAA12011.1; -.
DR EMBL; D83540; BAA12010.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR000461; Alpha-amylase.
DR InterPro; IPR002044; CBD_4.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF00686; CBD_4; 1.
DR ProDom; PD001568; CBD_4; 1.
DR Signal.
FT CHAIN 1 20 ALPHA-AMYLASE.
FT SIGNAL 21 631
SQ SEQUENCE 631 AA; 67658 MW; 819687B6E1D707E5 CRC64;
```

Query Match 75.0%; Score 33; DB 3; Length 631;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 NAFKSTSGS 9
DB 295 NAFKSTSGS 303

RESULT 11
ID Q917J5 PRELIMINARY; PRT; 688 AA.
AC Q917J5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIN PROTEIN.
GN RIN OR CG9412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V., Chapple M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fouts D.C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostali D., Houston K.A., Howland T.J., Wei M.-H., Iideyam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko F., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefel F., Shen H.,
RA Shue B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003701; AAG22151.1; -.
DR FlyBase; FBgn0015778; rin.
DR InterPro; IPR002075; NTF2.
DR InterPro; IPR000504; RRM.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PSS0102; RRM; 1.
SQ SEQUENCE 688 AA; 74713 MW; 9E329949E7248C19 CRC64;
```

Query Match 75.0%; Score 33; DB 5; Length 688;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
| | | | | | |
Db 385 NLFKSTSSS 393

RESULT 12
Q9NH72 PRELIMINARY; PRT: 690 AA.

AC Q9NH72: 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE RASPUTIN.
GN RIN OR CG9412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20191744; PubMed=10725247;
RA Pazman C., Mayes C.A., Panto M., Haynes S.R., Mlodzik M.;
RT "Rasputin, the Drosophila homolog of the RasGAP SH3 binding protein,
functions in Ras and Rho mediated signaling.";
RL Development 127:1715-1725(2000).
DR EMBL: AF211031; AAF68949.1; -;
DR FlyBase: FBgn0015778; rin.
DR InterPro: IPR002075; NTF2.
DR InterPro: IPR000504; RRM.
DR Pfam: PF00076; rim; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
SQ SEQUENCE 690 AA; 74985 MW; 40C5BD7E9F5FA253 CRC64;

Query Match 75.0%; Score 33; DB 5; Length 690;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
| | | | | | |
Db 387 NLFKSTSSS 395

RESULT 13
Q9VFT4 PRELIMINARY; PRT: 690 AA.

AC Q9VFT4: 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CG9412 PROTEIN (LID31194P).
GN RIN OR CG9412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champs M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Hejman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., Moleod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guan H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBD databases.
DR EMBL: AE003701; AAF54965.1; -;
DR EMBL: AY058617; AAL13846.1; -;
DR FlyBase: FBgn0015778; rin.
DR InterPro: IPR002075; NTF2.
DR InterPro: IPR000504; RRM.
DR Pfam: PF02136; NTF2; 1.
DR Pfam: PF00076; rim; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
SQ SEQUENCE 690 AA; 74943 MW; 40C5AC6F9FAEB353 CRC64;

Query Match 75.0%; Score 33; DB 5; Length 690;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
| | | | | | |
Db 387 NLFKSTSSS 395

RESULT 14
Q9YAT6 PRELIMINARY; PRT: 100 AA.

AC Q9YAT6: 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE HYPOHETICAL 10.4 KDA PROTEIN APE1858.
GN APE1858.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcates; Desulfurococcaceae;
OC Aeropyrum.
NCBI_TaxID=56636;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1:
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL: AP000062; BAA80862.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 100 AA; 10356 MW; 6C17D0ED584681EA CRC64;

Query Match 72.7%; Score 32; DB 17; Length 100;
 Best Local Similarity 77.8%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
 ||| 1:111
 DB 43 NAFCSSTSGS 51

RESULT 15
 ID Q9EOM8 PRELIMINARY; PRT; 196 AA.
 AC Q9EOM8;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE NUCLEOPORIN 98 (FRAGMENT).
 GN NUP98.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wu X., van Deursen J.M.A.;
 RT "Alternative splicing of mouse Nup98 mRNA.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF201386; AAC42323.1; -.
 DR MGD: MGI:109404; Nup98.
 KW Porin.
 FT NON_TER 1
 FT NON_TER 196
 SQ SEQUENCE 196 AA; 19483 MW; 7F508FBD4D5AE3D3 CRC64;

Query Match 72.7%; Score 32; DB 11; Length 196;
 Best Local Similarity 77.8%; Pred. No. 63;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
 | 1 11111
 DB 125 NFGSTSGS 133

Search completed: June 13, 2002, 09:11:15
 Job time: 1060 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:53 ; Search time 38.1 Seconds
(without alignments)
9.146 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Perfect score: 44
Sequence: 1 NAFKSTGS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	44	100.0	AMYA_ASPAW	Q02905 aspergillus
2	44	100.0	AMYA_ASPOR	P10529 aspergillus
3	44	100.0	AMYE_ASPAW	Q02906 aspergillus
4	44	100.0	AMY_ASPSH	P30292 aspergillus
5	32	72.7	FLG6_BORBU	O51715 borrelia bu
6	32	72.7	SLAF_AERSA	P35923 aeromonas s
7	32	72.7	NU98_HUMAN	P52948 homo sapien
8	31	70.5	ZRF1_MOUSE	P54103 mus musculu
9	31	70.5	CID_DROME	P15538 drosophila
10	30	68.2	TPMT_MOUSE	O55060 mus musculu
11	30	68.2	TPMT_MUSSP	O99X22 mus spreus
12	30	68.2	RUBA_RICCN	Q94187 rickettsia
13	30	68.2	EL3C_TOBAC	P23432 nicotiana t
14	30	68.2	DPAL_SALTY	P40817 arabidopsis
15	30	68.2	FABR_ARATH	P56271 aspergillus
16	30	68.2	AMYA_ASPNG	O95440 helicobacte
17	30	68.2	FLHA_HELPJ	O06758 helicobacte
18	30	68.2	FLHA_HELPJ	O06758 helicobacte
19	30	68.2	APAF_BRARE	Q91998 brachydanio
20	29	65.9	RBS1_RHOSH	P27998 rhodobacter
21	29	65.9	VG14_BPPH2	P11188 bacteriopho
22	29	65.9	VG14_BPPH2	P07539 bacteriopho
23	29	65.9	CYPH_ECHGR	P14088 streptomyce
24	29	65.9	PPI_STRCH	O06118 streptomyce
25	29	65.9	YT22_CAEEL	O10932 caenorhabdi
26	29	65.9	MOVV_AMYLE	P03595 alfalfa mos
27	29	65.9	MOVV_AMYLE	P03672 alfalfa mos
28	29	65.9	MOVV_AMYST	P03596 alfalfa mos
29	29	65.9	MOVV_AMYST	P24265 alfalfa mos
30	29	65.9	Y306_MYCPN	P73343 mycoplasma
31	29	65.9	ODO2_YEAST	P19262 saccharomyc
32	29	65.9	RK11_SECCE	O02723 secale cere
33	29	65.9	RRPP_CDVO	P06940 canine dist

34	29	65.9	710	1	IRAI_MOUSE	Q62406 mus musculu
35	29	65.9	759	1	RRP3_IANN	P21428 influenza a
36	29	65.9	759	1	RRP3_IABUD	P26104 influenza a
37	29	65.9	759	1	RRP3_INCHI	P11486 influenza a
38	29	65.9	759	1	RRP3_IAPOM	O82573 influenza a
39	29	65.9	759	1	RRP3_IAPPR	P12445 influenza a
40	29	65.9	759	1	RRP3_IAGU2	P18883 influenza a
41	29	65.9	759	1	RRP3_IAGU2	P26115 influenza a
42	29	65.9	759	1	RRP3_IAGU2	P26116 influenza a
43	29	65.9	759	1	RRP3_IAGU2	P26105 influenza a
44	29	65.9	759	1	RRP3_IAGU2	P26106 influenza a
45	29	65.9	759	1	RRP3_IAGU2	P16884 influenza a

ALIGNMENTS

RESULT	ID	AMYA_ASPAW	STANDARD	PRT	498 AA.
AC	Q02905				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan				
DE	glucanohydrolase A).				
GN	AMYA.				
OS	Aspergillus awamori.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_TaxID=105351;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=UVR143F;				
RX	MEDLINE=90254827; PubMed=2340591;				
RA	Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,				
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,				
RA	Berka R.M.;				
RT	"Cloning, characterization, and expression of two alpha-amylase genes				
RT	from Aspergillus niger var. awamori.";				
RL	Curr. Genet. 17:203-212(1990).				
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic				
CC	linkages in oligosaccharides and polysaccharides.				
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO				
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: X52755; CAA36966.1; -				
DR	HSSP: P10529; 7TAA.				
DR	InterPro: IPR000461; Alpha-amylase.				
DR	Pfam: PF00128; alpha-amylase; 1.				
KW	Glycoprotein; Signal; Multigene family.				
FT	SIGNAL	1	21	POTENTIAL.	
FT	CHAIN	22	498	ALPHA-AMYLASE A.	
FT	ACT_SITE	227	227	BY SIMILARITY.	
FT	ACT_SITE	231	231	BY SIMILARITY.	
FT	ACT_SITE	251	251	BY SIMILARITY.	
FT	ACT_SITE	318	318	BY SIMILARITY.	
FT	DISULFID	51	59	BY SIMILARITY.	
FT	DISULFID	171	185	BY SIMILARITY.	
FT	DISULFID	261	304	BY SIMILARITY.	
FT	DISULFID	461	496	BY SIMILARITY.	
FT	CARBOHYD	218	218	N-LINKED (GLCNAC...) (POTENTIAL).	
SO	SEQUENCE	498 AA;	54880 MW;	765851 IBC01A8A.01 CRC64;	

```

Query Match          100.0%; Score 44; DB 1; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.088;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NAFKSTSGS 9
        |||||||
DB      281 NAFKSTSGS 289

RESULT 2
AMTA_ASPOR STANDARD: PRT; 499 AA.
ID AMYA_ASPOR 010529; P11763; Q00250;
AC 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
DE (1,4-alpha-D-glucan glucoamylolase).
GN AMY1 AND AMY2 AND AMY3.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 63303;
RA MEDLINE=89237897; PubMed=2785629;
RT Wirel S., Lachmund A., Wildhardt G., Rutkowski E.;
RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
RT intron-exon organization.";
RL Mol. Microbiol. 3:3-14(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=89378767; PubMed=2789162;
RX Genes M.J., Dove M.J., Seligy V.L.;
RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
RT containing eight introns.";
RL Gene 79:107-117(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
RT of Aspergillus oryzae.";
RL Agric. Biol. Chem. 53:593-599(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=90128276; PubMed=2612911;
RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
RA Udaoka S.;
RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
RT evidence for multiple related genes.";
RL Gene 84:319-327(1989).
RN [5]
RP SEQUENCE OF 22-499.
RA Toda H., Kondo K., Narita K.;
RT "The complete amino acid sequence of Taka-amylase A.";
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
RN [6]
RP SEQUENCE OF 206-225.
RX MEDLINE=74001521; PubMed=4733850;
RA Isemura S., Ikenaka T.;
RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
RT A with trypsin and chymotrypsin.";
RL J. Biochem. 74:1-10(1973).
RN [7]
RP SEQUENCE OF 433-499.
RA Narita K.;
RT Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX MEDLINE=80227691; PubMed=6156152;

```

```

RA Matsunura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasunaka N.,
RA Toda H., Narita K., Kakudo M.;
RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
RT 3-A resolution.";
RL J. Biochem. 87:1555-1558(1980).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX MEDLINE=84212370; PubMed=6609921;
RA Matsunura Y., Kusunoki M., Harada W., Kakudo M.;
RT "Structure and possible catalytic residues of Taka-amylase A.";
RL J. Biochem. 95:697-702(1984).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
RX MEDLINE=97428212; PubMed=9283074;
RA Brzozowski A.M., Davies G.J.;
RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
RT inhibitor acarbose at 2.0-A resolution.";
RL Biochemistry 36:10837-10845(1997).
CC -|- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -|- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -|- MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
CC -|- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12725; CAA31218.1; -
DR EMBL; X12726; CAA31219.1; -
DR EMBL; X12727; CAA31220.1; -
DR EMBL; D00434; BAA00336.1; -
DR EMBL; M33218; AAA32708.1; -
DR PIR; S04548; ALAS1.
DR PIR; S04549; ALAS3.
DR PIR; A10627; A10627.
DR PIR; A10627; A10627.
DR PIR; JK0201; JK0201.
DR PIR; JS0240; JS0240.
DR PIR; JT0466; JT0466.
DR PDB; 2TAA; 15-OCT-89.
DR PDB; 6TAA; 31-OCT-93.
DR PDB; 7TAA; 25-NOV-98.
DR GLYCOSULEDB; P10529; -.
DR InterPro; IPR000461; Alpha-amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family; 3D-structure.
FT SIGNAL 1 21
FT CHAIN 22 499
FT ACT_SITE 227 227
FT ACT_SITE 231 231
FT ACT_SITE 251 251
FT ACT_SITE 318 318
FT DISULFID 51 39
FT DISULFID 171 185
FT DISULFID 261 304
FT DISULFID 461 496
FT CARBOHYD 218 218
FT VARIANT 56 56
FT VARIANT 172 172
FT VARIANT 93 94
FT CONFLICT 106 106
FT CONFLICT 184 184
FT CONFLICT 195 195
FT CONFLICT 255 255
FT CONFLICT 291 291
FT CONFLICT 345 345
FT N-LINKED (GLCNAC. . .).
FT /FTid=CAR.000125.
FT Q -> R (IN AMY3).
FT TT -> L (IN AMY3).
FT TT -> DC (IN REF. 5).
FT Q -> T (IN REF. 5).
FT D -> Y (IN REF. 3).
FT P -> L (IN REF. 3).
FT G -> V (IN REF. 3).
FT D -> H (IN REF. 4).
FT I -> L (IN REF. 5).

```

```
FT CONFLICT 370 370 L -> A (IN REF. 4).
FT CONFLICT 406 409 WEIY -> PYI (IN REF. 5).
FT CONFLICT 448 448 G -> S (IN REF. 5).
FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).
FT HELIX 24 27
FT TURN 28 29
FT STRAND 32 35
FT HELIX 37 40
FT STRAND 42 42
FT TURN 43 44
FT HELIX 53 55
FT STRAND 61 61
FT HELIX 63 67
FT TURN 68 69
FT HELIX 70 74
FT TURN 75 77
FT STRAND 80 83
FT STRAND 87 89
FT STRAND 94 94
FT TURN 95 96
FT STRAND 97 97
FT TURN 100 101
FT STRAND 105 111
FT TURN 113 115
FT TURN 118 119
FT HELIX 120 130
FT TURN 131 132
FT STRAND 134 139
FT STRAND 143 143
FT STRAND 146 146
FT HELIX 150 152
FT TURN 155 156
FT STRAND 159 159
FT HELIX 164 166
FT STRAND 167 167
FT STRAND 172 172
FT HELIX 179 184
FT STRAND 186 188
FT STRAND 193 194
FT STRAND 196 197
FT TURN 199 200
FT HELIX 202 219
FT TURN 220 220
FT STRAND 223 226
FT TURN 227 228
FT HELIX 229 231
FT TURN 234 235
FT HELIX 236 244
FT STRAND 247 250
FT HELIX 257 265
FT TURN 266 266
FT STRAND 269 271
FT HELIX 273 283
FT TURN 286 287
FT HELIX 290 303
FT STRAND 307 309
FT TURN 311 312
FT TURN 317 318
FT HELIX 322 325
FT HELIX 329 341
FT STRAND 345 349
FT TURN 350 351
FT HELIX 352 354
FT TURN 355 355
FT TURN 361 364
FT HELIX 368 371
FT TURN 372 372

Query Match
Best Local Similarity 100.0%; Score 44; DB 1; Length 499;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
DB 281 NAFKSTSGS 289
|||||
RESULT 3
ID AMYB ASPAW STANDARD; PRT; 499 AA.
AC 002906:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha-amyrase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase B).
GN AMYB.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UW143F;
RX MEDLINE=90234827; Pubmed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amyase genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52756; CAA36967.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR000461; Alpha_amyase.
DR Pfam; PF00128; alpha-amyase.1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 499
FT ACT_SITE 227 227 BY SIMILARITY.
FT ACT_SITE 231 231 BY SIMILARITY.
FT ACT_SITE 251 251 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 261 304 BY SIMILARITY.
FT DISULFID 461 496 BY SIMILARITY.
FT CARBOHYD 218 218 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 499 AA; 54921 MW; 740B96B1BEC01A8A CRC64;

Query Match
Best Local Similarity 100.0%; Score 44; DB 1; Length 499;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 4
AMY ASPSH

```

ID  AMY_ASPEH  STANDARD;  PRT;  499 AA.
AC  P30292;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amyase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY.
OS  Aspergillus shirousami.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC  Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX  NCBI_TaxID=5070;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92323146; PubMed=1368777;
RA  Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT  "Cloning of the alpha-amyase cDNA of Aspergillus shirousami and its
RT  expression in Saccharomyces cerevisiae."
RL  Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC  -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; D10461; BA01255.1; -.
DR  PIR; J50663; J50663.
DR  HSSP; P10529; 7TAA.
DR  InterPro; IPR000461; Alpha-amyase.
DR  Pfam; PF00128; alpha-amyase; 1.
KM  Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KM  Glycoprotein; Signal.
FT  SIGNAL 1 21
FT  CHAIN 22 499 ALPHA-AMYLASE.
FT  ACT_SITE 227 227
FT  ACT_SITE 231 231 BY SIMILARITY.
FT  ACT_SITE 251 251 BY SIMILARITY.
FT  ACT_SITE 318 318 BY SIMILARITY.
FT  DISULFID 51 59 BY SIMILARITY.
FT  DISULFID 171 185 BY SIMILARITY.
FT  DISULFID 261 304 BY SIMILARITY.
FT  DISULFID 461 496 BY SIMILARITY.
FT  CARBOHD 218 218 N-LINKED (GLCNAc. . .) (BY SIMILARITY).
SQ  SEQUENCE 499 AA; 54852 MW; 1FB7AE50DA01C03F CRC64;

Query Match 100.0%; Score 44; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.088;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
DB 281 NAFKSTSGS 289

RESULT 5
FLUG_BORBU STANDARD; PRT; 265 AA.
ID FLUG_BORBU
AC 051715;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar basal-body rod protein flug (Distal rod protein).
GN FLUG OR BB0774.
OS Borrelia burgdorferi (Lyme disease spirochete).

```

```

OC  Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX  NCBI_TaxID=139;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 35210 / B31;
RX  MEDLINE=98065943; PubMed=9403685;
RA  Fraser C.M., Castens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA  Lathiga R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA  Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA  Peterson J., Kellavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA  van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA  Uterback T., Matthey L., McDonald L., Arlrich P., Bowman C.,
RA  Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA  Smith H.O., Venter J.C.;
RT  "Genomic sequence of a Lyme disease spirochete, Borrelia
RT  burgdorferi".
RL  Nature 390:580-586(1997).
CC  -1- SUBUNIT: THE BASAL BODY CONSTITUTES A MAJOR PORTION OF THE
CC  FLAGELLAR ORGANELLE AND CONSISTS OF FOUR RINGS (L,P,S, AND M)
CC  MOUNTED ON A CENTRAL ROD. THE ROD CONSISTS OF ABOUT 26 SUBUNITS OF
CC  FLAG IN THE DISTAL PORTION, AND FLGB, FLGC AND FLGF ARE THOUGHT TO
CC  BUILD UP THE PROXIMAL PORTION OF THE ROD WITH ABOUT 6 SUBUNITS
CC  EACH (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE FLAGELLA BASAL BODY ROD PROTEINS
CC  FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE001177; AAC67132.1; -.
DR  TIGR; BB0774; -.
DR  InterPro; IPR001444; Flag_ba-rod.
DR  Pfam; PF00460; flg_ba-rod; 1.
DR  PROSITE; PS00588; FLAGELLA_BB_ROD; 1.
KM  Flagella; Complete proteome.
SQ  SEQUENCE 265 AA; 28960 MW; EEA5E79B8FCBDBD CRC64;

Query Match 72.7%; Score 32; DB 1; Length 265;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
DB 192 NLFKSTSGS 200

RESULT 6
SIAP_AERSA STANDARD; PRT; 502 AA.
ID SIAP_AERSA
AC  P35823;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE S-layer protein precursor (Tetragonal surface virulence array
DE  protein) (A-protein).
GN  VAPA.
OS  Aeromonas salmonicida.
OC  Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC  Aeromonas.
OX  NCBI_TaxID=645;
RN  [1]
RP  SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC  STRAIN=A450;
RX  MEDLINE=91332048; PubMed=1869553;
RA  Chu S., Cavaignac S., Feutrier J., Phipps B.M., Kostorzynska M.,
RA  Kay W.W., Trust T.J.;
RT  "Structure of the tetragonal surface virulence array protein and gene

```

```

RT of Aeromonas salmonicida."
RN J. Biol. Chem. 266:15258-15265(1991).
RP SEQUENCE OF 22-48.
RC STRAIN-A400, A450, AND A461:
RX MEDLINE=85124084; PubMed=6525563;
RA Kay W.W., Phipps B.M., Ishiguro E.E., Olafson R.W., Trust T.J.;
RT "Surface layer virulence A-proteins from Aeromonas salmonicida
strains."
RL Can. J. Biochem. Cell Biol. 62:1064-1071(1984).
RN [13]
RP SEQUENCE OF 22-48.
RC STRAIN-A450;
RX MEDLINE=83257214; PubMed=6871174;
RA Phipps B.M., Trust T.J., Ishiguro E.E., Kay W.W.;
RT "Purification and characterization of the cell surface virulent A
protein from Aeromonas salmonicida."
RL Biochemistry 22:2934-2939(1983).
RN [4]
RP BINDING PROPERTIES.
RX MEDLINE=92112689; PubMed=1730607;
RA Dohg P., Emoedy L., Trust T.J.;
RT "Binding of laminin and fibronectin by the trypsin-resistant major
structural domain of the crystalline virulence surface array protein
of Aeromonas salmonicida."
RL J. Biol. Chem. 267:43-49(1992).
CC -1- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. THIS S-LAYER
CONFERES VIRULENCE ON THE BACTERIUM INFECTING SALMONID AND NON
SALMONID FISH. IT BINDS SPECIFICALLY TO LAMININ AND FIBRONECTIN.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
S-LAYER WITH TETRAGONAL SYMMETRY. TETHERED TO THE OUTER MEMBRANE
BY INTERACTION WITH O-SIDE CHAIN OF THE LPS.
CC -1- DOMAIN: THIS S-LAYER PROTEIN IS COMPOSED OF TWO DISTINCT
STRUCTURAL DOMAINS: A LARGER MASS TRYPSIN-RESISTANT N-TERMINAL,
CONTAINING THE BINDING REGION FOR EXTRACELLULAR MATRIX PROTEINS,
AND A SMALLER C-TERMINAL DOMAIN WITH INTERMEDIATE RESISTANCE TO
TRYPSIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M64655; AAA21941.1; -.
KW Signal; Cell wall; S-layer; Virulence.
FT SIGNAL 1 21
FT CHAIN 22 502
FT VARIANT 26 26 S->G (IN STRAINS A450 AND A461).
FT VARIANT 35 35 T->N (IN STRAIN A450).
FT VARIANT 47 47 D->S (IN STRAINS A400 AND A450).
FT VARIANT 48 48 F->D (IN STRAIN A400).
SQ SEQUENCE 502 AA; 52868 MW; 8B86AC6C4FE46F77 CRC64;

Query Match 72.7%; Score 32; DB 1; Length 502;
Best Local Similarity 87.5%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AFKSTSGS 9
DB 176 AFKLTSGS 183

RESULT 7
NU98_HUMAN STANDARD: PRT; 920 AA.
AC P52948;
DT 01-OCT-1996 (Rel. 34, Created)

```

```

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Nuclear pore complex protein Nup98 (Nucleoporin Nup98) (98 kDa
DE nucleoporin).
GN NUP98.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96154188; PubMed=8563754;
RA Borrow J., Shearman A.M., Stanton V.P., Becher R., Collins T.,
RA Williams A.J., Dube I., Katz F., Kwong Y.L., Morris C.,
RA Ohyashiki K., Toyama K., Rowley J., Housman D.E.;
RT "The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses
the genes for nucleoporin NUP98 and class I homeoprotein HOXA9."
RL Nat. Genet. 12:159-167(1996).
CC -1- FUNCTION: PLAYS A ROLE IN THE BIDIRECTIONAL TRANSPORT ACROSS THE
CC NUCLEOPORIN COMPLEX (NPC). THE REPEAT DOMAIN HAS A DIRECT ROLE IN
CC THE TRANSPORT.
CC -1- SUBCELLULAR LOCATION: Nuclear pore complex.
CC -1- DOMAIN: CONTAINS G-L-F-G REPEATS.
CC -1- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA IS CHARACTERIZED BY A
CC CHROMOSOMAL TRANSLOCATION T(7;11)(P15;P15) WHICH INVOLVES HOXA9
CC AND NUCLEOPORIN NUP98 GENES.
CC -1- SIMILARITY: BELONGS TO THE GLFG FAMILY OF NUCLEOPORINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U41815; AAC50366.1; -.
DR MIM; 601021; -.
DR InterPro: IPR004325; Nucleoporin_FG.
DR Pfam: PF03093; Nucleoporin_FG; 25.
KW Nuclear protein; Transport; Repeat; Chromosomal translocation.
FT DOMAIN 873 877 POLY-GLU.
SQ SEQUENCE 920 AA; 96074 MW; 129DABAFB6923AD CRC64;

Query Match 72.7%; Score 32; DB 1; Length 920;
Best Local Similarity 77.8%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 NAFKSTSGS 9
DB 138 NPFKSTSGS 146

RESULT 8
ZRF1_MOUSE STANDARD: PRT; 514 AA.
AC P54103;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zuo1in related factor-1.
GN ZRF1 OR DNAC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96115610; PubMed=8666407;
RA Hughes R., Chan F.Y., White R.A., Zon L.I.;
RT "Cloning and chromosomal localization of a mouse cDNA with homology
to the Saccharomyces cerevisiae gene zuo1in."

```

```

CC   Genomics 29:546-550(1995).
CC   -1- SUBCELLULAR LOCATION: Nuclear.
CC   -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES.
CC   -1- SIMILARITY: CONTAINS 1 J DOMAIN.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: U53208; AAC52486.1; -.
DR   HSSP: P25685; 1HDJ.
DR   MGD: MGI:9470; DnaJc2.
DR   InterPro: IPR001623; DnaJ_N.
DR   Pfam: PF00226; DnaJ_1.
DR   Pfam: PF00249; myb_DNA-binding; 1.
DR   SMART: SM00271; DnaJ; 1.
DR   SMART: SM00393; SANT; 1.
DR   PROSITE: PS00636; DnaJ_1; 1.
DR   PROSITE: PS50076; DnaJ_2; 1.
DR   Chapterone: Nuclear Protein.
KW   DOMAIN.
FT   DOMAIN 94..163 J-DOMAIN.
SQ   SEQUENCE 514 AA; 59510 MW; 10BFD5DFACFB9DE CRC64;
OY   1 NAFKSTGS 9
Db   439 NAFKSTGS 447
Query Match          70.5%; Score 31; DB 1; Length 514;
Best Local Similarity 77.8%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

RESULT 9

CID_DROME STANDARD; PRT: 1377 AA.

AC P19538;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE Cubitus interruptus dominant protein.

GN Ci-D.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

DE [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OREGON-R;

RC MEDLINE=90346286; PubMed=2166702;

RA Orenic T.V., Slusarski D.C., Kroll K.L., Holmgren R.A.;

RT "Cloning and characterization of the segment polarity gene cubitus interruptus dominant of drosophila.",

RL Genes Dev. 4:1053-1067(1990).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=OREGON-R;

RC MEDLINE=92146935; PubMed=1686006;

RA Berry A.J., Ajick J.W., Kreitman M.;

RT "Lack of polymorphism on the Drosophila fourth chromosome resulting from selection."

RT Genetics 129:1111-1117(1991).

CC -1- FUNCTION: INVOLVED IN SEGMENT POLARITY. IS REQUIRED FOR THE NORMAL DEVELOPMENT OF THE POSTERIOR HALF OF EACH EMBRYONIC SEGMENT.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED UNIFORMLY THROUGHOUT THE BLASTODERM STAGE AND GASTRULATION AND DOES NOT RESOLVE INTO SEGMENTALLY

```

CC   REPEATING STRIPES UNTIL THE END OF THE SHORT PHASE OF GERM-BAND
CC   EXTENSION.
CC   -1- SIMILARITY: TO THE GLI-RELATED GROUP OF C2H2-TYPE ZINC-FINGERS
CC   PROTEINS.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: X54360; CAA8244.1; -.
DR   PIR: A35817; A35817.
DR   PIR: S12769; S12769.
DR   HSSP: P08151; 2GLI.
DR   FlyBase: FBgn0004859; ci.
DR   InterPro: IPR000822; Znf-C2H2.
DR   Pfam: PF00096; Zf-C2H2; 5.
DR   SMART: PR00048; ZINC-FINGER.
DR   PRINTS: PR00355; Znf_C2H2; 5.
DR   PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE: PS50157; ZINC_FINGER_C2H2_2; 4.
KW Developmental protein; segmentation polarity protein; zinc-finger;
KW Metal-binding; DNA-binding; Repeat; Nuclear protein.
FT   DOMAIN 451..603 ZINC_FINGERS.
FT   ZN_FING 451..476 C2H2-TYPE.
FT   ZN_FING 484..511 C2H2-TYPE.
FT   ZN_FING 517..541 C2H2-TYPE.
FT   ZN_FING 547..572 C2H2-TYPE.
FT   ZN_FING 578..603 C2H2-TYPE.
SQ   SEQUENCE 1377 AA; 150881 MW; 85189314D5E7B3B7 CRC64;
OY   1 NAFKSTGS 8
Db   901 NSFSTSG 908
Query Match          70.5%; Score 31; DB 1; Length 1377;
Best Local Similarity 75.0%; Pred. No. 14e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

RESULT 10

TPMT_MOUSE STANDARD; PRT: 240 AA.

AC O55060; O90UG7; O9JIL7;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Thiopurine S-methyltransferase (EC 2.1.1.67) (Thiopurine methyltransferase).

GN TPMT.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C3H;

RC MEDLINE=98198541; PubMed=9539138;

RA Fessing M.Y., Belkov V.M., Krynetski E.Y., Evans W.E.;

RT "Molecular cloning and functional characterization of the cDNA encoding the murine thiopurine S-methyltransferase (TPMT)."

RL FBS Lett. 424:143-145(1998).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=DDA/J2J, AND C57BL/6J;

RA Adjei A.A., Johnson G.B., O'Leary D.M., Weinshilboum R.M.;

RT "Mouse thiopurine methyltransferase pharmacogenetics: cDNA cloning and characterization and processed pseudogene cloning."

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.


```

RN [3]
RP SEQUENCE OF 43-227 FROM N.A.
RC STRAIN=129/OLA;
RL Krynetski E.Y., Fessing M.Y., Edick M.J., Evans W.E.;
RA Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE S-METHYLATION OF THIOPURINE DRUGS SUCH AS
CC 6-MERCAPTOPURINE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + a thiopurine = S-
CC adenosyl-L-homocysteine + a thiopurine S-methylether.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. TPMT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF046887; AAC25919.1; -
DR EMBL; AF037043; AAD02092.1; -
DR EMBL; AF037044; AAD02093.1; -
DR EMBL; AF104832; AAF06075.1; -
DR EMBL; AF104825; AAF06075.1; JOINED.
DR EMBL; AF104826; AAF06075.1; JOINED.
DR EMBL; AF104827; AAF06075.1; JOINED.
DR EMBL; AF104828; AAF06075.1; JOINED.
DR EMBL; AF104829; AAF06075.1; JOINED.
DR EMBL; AF104830; AAF06075.1; JOINED.
DR EMBL; AF104831; AAF06075.1; JOINED.
DR EMBL; AF218593; AAF74424.1; -
DR EMBL; AF218588; AAF74424.1; JOINED.
DR EMBL; AF218589; AAF74424.1; JOINED.
DR EMBL; AF218590; AAF74424.1; JOINED.
DR EMBL; AF218591; AAF74424.1; JOINED.
DR EMBL; AF218592; AAF74424.1; JOINED.
DR MGD; MGI:98812; Tpmf.
RN Transferase; Methyltransferase.
FT VARIANT 69 69 I -> V (IN STRAIN C57BL/6J).
SQ SEQUENCE 240 AA; 27585 MW; 2BA57F30E8EB72D2 CRC64;

```

```

Query Match 68.2%; Score 30; DB 1; Length 240;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 FKSTSGS 9
DB 116 FKSSSGS 122

```

```

RESULT 11
TPMT_MUSP STANDARD; PRT; 240 AA.
AC 090X22;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Thiopurine S-methyltransferase (EC 2.1.1.67) (Thiopurine
DE methyltransferase).
GN TPMT.
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBT_TaxID=10096;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=EI;
RA Adlei A.A., Johnson G.B., Otterness D.M., Weinsilbom R.M.;
RT "Mouse thiopurine methyltransferase pharmacogenetics: cDNA cloning and

```

```

RT characterization and processed pseudogene cloning."
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE S-METHYLATION OF THIOPURINE DRUGS SUCH AS
CC 6-MERCAPTOPURINE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + a thiopurine = S-
CC adenosyl-L-homocysteine + a thiopurine S-methylether.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. TPMT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF037045; AAD02094.1; -
DR MGD; MGI:98812; Tpmf.
RN Transferase; Methyltransferase.
SQ SEQUENCE 240 AA; 27617 MW; 860A0DC5BD6FBA4C CRC64;

```

```

Query Match 68.2%; Score 30; DB 1; Length 240;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 FKSTSGS 9
DB 116 FKSSSGS 122

```

```

RESULT 12
RUVB_RICCN STANDARD; PRT; 342 AA.
AC 092I87;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Holliday junction DNA helicase ruvb.
GN RUVB OR RC0533.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBT_TaxID=781;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RT Science 293:2093-2098(2001).
CC -1- FUNCTION: The ruvb-ruvb complex in the presence of ATP renatures
CC cruciform structure in supercoiled DNA with palindromic sequence,
CC indicating that it may promote strand exchange reactions in
CC homologous recombination. RuvbA is an helicase that mediates the
CC Holliday junction migration by localized denaturation and
CC reannealing (By similarity).
CC -1- SUBUNIT: Forms a complex with ruvA (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RUVB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: AE008615; AAL03071.1; ALT_INIT.
 DR InterPro: IPR003959; AAA_ATPase_cent.
 DR InterPro: IPR002182; NB-ARC.
 DR Pfam: PF00904; AAA.1.
 DR Pfam: PF00931; NB-ARC.1.
 DR DNA repair: SOS response; ATP-binding; DNA recombination; Helicase;
 KW Complete proteome.
 FT NP_BIND 57 64 ATP (POTENTIAL).
 SQ SEQUENCE 342 AA; 38309 MW; 83891B80C829C28 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FRSTSG 8
 |||||
 DB 78 FRSTSG 83

RESULT 13
 E13C_TOBAC STANDARD: PRT; 351 AA.

AC P23432;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Glucan endo-1,3-beta-glucosidase precursor (EC 3.2.1.39) ((1->3)-beta-glucan endohydrolase) ((1->3)-beta-glucanase) (Beta-1,3-endoglucanase).
 DE SP41A.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE-91006039; PubMed-2120041;
 RA Ori N., Sessa G., Lotan T., Himmelhoch S., Fluhr R.;
 RT "A major stylar matrix polypeptide (sp41) is a member of the pathogenesis-related proteins superclass.";
 RL EMBO J. 9:3429-3436(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-CV. SAMSUN NN. TISSUE-Style;
 RX MEDLINE-96145508; PubMed-8555460;
 RA Sessa G., Fluhr R.;
 RT "The expression of an abundant transmitting tract-specific endoglucanase (Sp41) is promoter-dependent and not essential for the reproductive physiology of tobacco.";
 RL Plant Mol. Biol. 29:969-982(1995).
 CC -1- FUNCTION: IMPLICATED IN THE DEFENSE OF PLANTS AGAINST PATHOGENS.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,3-beta-D-glucosidic linkages in 1,3-beta-D-glucans.
 CC -1- SUBCELLULAR LOCATION: STYLAR SECRETORY MATRIX.
 CC -1- DEVELOPMENTAL STAGE: MAXIMAL LEVEL OF ACCUMULATION DURING ANTHESIS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 17 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X54430; CAA38302.1; -
 DR EMBL: X81560; CAA57255.1; -
 DR PIR: S12013; S12013.
 DR HSSP: P15737; IGHS.
 DR InterPro: IPR000490; Glyco_hydro_17.

DR Pfam: PF00332; Glyco_hydro_17; 1.
 DR PROSITE: PS00587; GLYCOSYL_HYDROL_F17; 1.
 KW Hydrolyase; Glycosidase; Glycoprotein; signal; Multigene family.
 FT STGNL 1 32
 FT CHAIN 33 351
 FT MOD_RES 33 33
 FT ACT_SITE 268 268
 FT ACT_SITE 328 328
 FT CARBOHYD 79 79
 FT CARBOHYD 99 99
 FT CARBOHYD 235 235
 FT CARBOHYD 337 337
 SQ SEQUENCE 351 AA; 39186 MW; 588B82C3E8C71F39 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 351;
 Best Local Similarity 66.7%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAKSTSGS 9
 |||||
 DB 73 NIKSLNGS 81

RESULT 14
 DPAL_SALTY STANDARD: PRT; 404 AA.

AC P40817;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Diaminopropionate ammonia-lyase (EC 4.3.1.15) (Diaminopropionatase)
 DE (Alpha,beta-diaminopropionate ammonia-lyase).
 GN DPAL OR STM1002.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-LP2 / SGSC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
 RA "Complete genome sequence of *Salmonella enterica* serovar Typhimurium LP2.";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE OF 1-22; 67-88 AND 239-242.
 RX STRAIN-IFO 12529;
 RX MEDLINE-88087224; PubMed-3275662;
 RA Nagasawa T., Tanizawa K., Satoda T., Yamada H.;
 RT "Diaminopropionate ammonia-lyase from *Salmonella typhimurium*. Purification and characterization of the crystalline enzyme, and sequence determination of the pyridoxal 5'-phosphate binding peptide.";
 RL J. Biol. Chem. 263:958-964(1988).
 CC -1- FUNCTION: CATALYZES THE ALPHA,BETA-ELIMINATION REACTION OF BOTH L- AND D-ALPHA,BETA-DIAMINOPROPIONATE, THE MOST SUITABLE SUBSTRATES TO FORM PYRUVATE AND AMMONIA. THE L- AND D-ISOMERS OF SERINE ARE ALSO DEGRADED, THOUGH SLOWLY; IT IS THE ONLY SERINE DEHYDRATASE WHICH CAN ELIMINATE AN AMINO GROUP AT THE BETA-CARBON POSITION.
 CC -1- CATALYTIC ACTIVITY: 2,3-diaminopropionate = pyruvate + 2 NH(3).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- ENZYME REGULATION: COMPETITIVELY INHIBITED BY L- AND D-ALANINE.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: TO E. COLI YGEX AND TO THREONINE DEHYDRATASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AE008743; AAL19936.1; -
 DR PIR: A28524; A28524.
 DR StyGene: SG10535; dpaL.
 KW Lyase: Pyridoxal phosphate: Complete proteome.
 FT BINDING 78 78 PYRIDOXAL PHOSPHATE (POTENTIAL).
 FT CONFLICT 16 16 Y -> K (IN REF. 2).
 FT CONFLICT 68 68 E -> G (IN REF. 2).
 FT CONFLICT 242 242 F -> A (IN REF. 2).
 SQ SEQUENCE 404 AA; 44152 MW; 9D605DFEA691F00 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 404;
 Best Local Similarity 66.7%; Pred. No. 63;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
 |||||
 Db 75 NAFKGLGGS 83

RESULT 15
 FABB_ARATH
 ID FABB_ARATH STANDARD; PRT; 473 AA.
 AC P52410;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 3-oxoacyl-[acyl-carrier-protein] synthase I, chloroplast precursor
 (EC 2.3.1.41) (Beta-ketoacyl-ACP synthase I) (KAS I).
 GN KAS1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID:3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Miller A.A., Kunst L.;
 RT "Isolation of an Arabidopsis cDNA encoding 3-ketoacyl-acyl carrier
 RT protein synthase I.1.";
 RL (In) Plant Gene Register PGR95-027.
 CC -1- FUNCTION: CATALYZES THE CONDENSATION REACTION OF FATTY ACID
 CC SYNTHESIS BY THE ADDITION TO AN ACYL ACCEPTOR OF TWO CARBONS
 CC FROM MALONYL-ACP. SPECIFIC FOR ELONGATION FROM C-10 TO UNSATURATED
 CC C-16 AND C-18 FATTY ACIDS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl]-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE BETA-KETOACYL-ACP SYNTHASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: U24177; AAC49118.1; -
 DR HSSP: P39435; IKAS.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR Pfam: PF00109; ketoacyl-synt; 1.
 DR Pfam: PF02801; ketoacyl-synt_C; 1.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.

KW Fatty acid biosynthesis: Transferase; Chloroplast; Transit peptide;
 KW Acyltransferase.
 FT TRANSIT 1 48 CHLOROPLAST (POTENTIAL).
 FT CHAIN 49 473 3-OXOACYL-[ACYL-CARRIER-PROTEIN] SYNTHASE
 FT ACT_SITE 224 224 BY SIMILARITY.
 SQ SEQUENCE 473 AA; 5185F91A1762008 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 473;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FKSTSG 8
 |||||
 Db 383 FKSTSG 388

Search completed: June 13, 2002, 09:11:54
 Job time: 1034 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:55:53 ; Search time 73.31 Seconds
(without alignments)
11.797 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Perfect score: 44

Sequence: 1 NAFKSTSGS 9

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	478	2 JK0201	alpha-amylase (EC
2	44	100.0	498	2 A48305	alpha-amylase (EC
3	44	100.0	499	1 ALAS1	alpha-amylase (EC
4	44	100.0	499	1 ALAS1	alpha-amylase (EC
5	44	100.0	499	2 B48305	alpha-amylase (EC
6	44	100.0	499	2 JS0663	alpha-amylase (EC
7	44	100.0	499	2 JT0466	alpha-amylase (EC
8	44	100.0	499	2 JN0588	alpha-amylase (EC
9	35	79.5	469	2 S17726	H+-transporting AT
10	34	77.3	102	2 S69741	hypothetical prote
11	33	75.0	297	2 H83888	transmembrane lipo
12	33	75.0	631	2 S72270	alpha-amylase (EC
13	32	72.7	100	2 A72572	hypothetical prote
14	32	72.7	265	2 E70196	flagellar basal bo
15	32	72.7	502	2 A39351	tetragonal surface
16	32	72.7	513	1 S60303	serine/threonine-s
17	32	72.7	681	2 T16353	hypothetical prote
18	32	72.7	1649	2 C86822	hypothetical prote
19	31	70.5	234	2 AG1512	hypothetical prote
20	31	70.5	251	2 AH2520	hypothetical prote
21	31	70.5	419	2 S41607	atrolysin A (EC 3.
22	31	70.5	471	2 B96531	hypothetical prote
23	31	70.5	492	2 F86868	chitinase (EC 3.2.
24	31	70.5	492	2 T47146	hypothetical prote
25	31	70.5	621	2 A57591	Id-associated prote
26	31	70.5	737	2 T02178	hypothetical prote
27	31	70.5	760	2 T41644	hypothetical trp-a
28	31	70.5	766	2 G71437	probable resistant
29	31	70.5	1041	2 C85189	disease resistance

30	31	70.5	1103	2 A85189	disease resistance
31	31	70.5	1125	2 H87644	Tom-dependent rec
32	31	70.5	1130	2 T21134	hypothetical prote
33	31	70.5	1184	2 H71436	hypothetical prote
34	31	70.5	1256	2 C71436	probable resistance
35	31	70.5	1301	2 D85188	disease resistance
36	31	70.5	1304	2 G85188	disease resistance
37	31	70.5	1317	2 B85189	disease resistance
38	31	70.5	1377	2 A38926	DNA-binding protei
39	31	70.5	1422	2 B71437	probable membrane
40	31	70.5	1484	2 C97196	probable resistance
41	31	70.5	1588	2 A86036	probable adhesin Z
42	31	70.5	1588	2 H91188	probable adhesin E
43	31	70.5	2088	2 E71436	hypothetical prote
44	31	70.5	2467	2 D71437	probable resistant
45	30	68.2	50	2 H70224	hypothetical prote

ALIGNMENTS

RESULT 1

JK0201

alpha-amylase (EC 3.2.1.1) - Aspergillus oryzae

N:Alternate names: glycogenase; Taka-amylase A

C:Species: Aspergillus oryzae

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998

C:Accession: JK0201

R:Toda, H.; Kondo, K.; Narita, K.

Proc. Jpn. Acad. 58B, 208-212, 1982

A:Title: The complete amino acid sequence of Taka-amylase A.

A:Reference number: JK0201

A:Molecule type: protein

A:Residues: 1478 <TOD>

C:Comment: One atom of calcium per molecule is essential for the activity.

C:Comment: This enzyme is a glycoprotein.

C:Function: See also PIR:JT0466 and PIR:JS0240.

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:173-300/Domain: alpha-amylase core homology <AMY>

F:197/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 44; DB 2; Length 478;

Best Local Similarity 100.0%; Pred. No. 0.19; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9

DB 260 NAFKSTSGS 268

RESULT 2

A48305

alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori

C:Species: Aspergillus awamori

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997

C:Accession: A48305

R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J

Cur. Genet. 17, 203-212, 1990

A:Title: Cloning, characterization, and expression of two alpha-amylase genes from As

A:Reference number: A48305

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1498 <KOR>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 44; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NAFKSTSGS 9
Db 281 NAFKSTSGS 289

RESULT 3
ALAS1
alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A

C:Species: Aspergillus oryzae
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2000

C:Accession: S04548; A33214; J50240; A91930; A93767; A10627
R:Witsell, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon
A:Reference number: S04548; MUID:89237897

A:Accession: S04548
A:Molecule type: DNA

A:Residues: 1-499 <WR>
A:Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921

A:Genetics: AMY1

A:Accession: A33214
A:Molecule type: mRNA

A:Residues: 1-499 <MT>
A:Cross-references: GB:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921

R:Genes, M.J.; Dove, M.J.; Seliy, V.L.
Gene 79, 107-117, 1989

A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
A:Reference number: J50240; MUID:89378767

A:Accession: J50240
A:Molecule type: DNA

A:Residues: 1-499 <GEN>
A:Genetics: AMY2

A>Note: the authors refer to this as isozyme II

R:Isemura, S.; Ikenaka, T.
J. Biochem. 74, 1-10, 1973

A:Reference number: A91930; MUID:74001521
A:Accession: A91930
A:Molecule type: protein

A:Residues: 206-225 <ISR>
R:Marita, K.
Proc. Jpn. Acad. 51, 285-290, 1975

A:Reference number: A93767
A:Accession: A93767
A:Molecule type: protein

A:Residues: 434-443, 446-447, 'Q', 449-458, 'GTTV', 459-464, 467-468, 'B', 470, 'B', 472-499 <NAR>
R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984

A:Title: Structure and possible catalytic residues of Taka-amylase A.
A:Reference number: A37454; MUID:84212370

C:Contents: annotation; X-ray crystallography, 3.0 angstroms
R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkemburg, J.P.; Wilkinson, A.
Submitted to the Brookhaven Protein Data Bank, August 1992

A:Reference number: A51548; PDB:6TRA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497
C:Comment: One atom of calcium per molecule is essential for activity.

C:Genetics: <AMY1>
A:Gene: amy1

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Genetics: <AMY2>

A:Gene: amy2; Amy1I
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po

F:1-21/Domain: signal sequence #status predicted <Sig>
F:22-499/Product: alpha-amylase 1 #status experimental <MAT>

F:194-321/Domain: alpha-amylase core homology <AMY>
F:51-59, 171-185, 261-304, 461-496/disulfide bonds: #status experimental

F:142, 183, 196, 231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:227, 251, 318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 44; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NAFKSTSGS 9
Db 281 NAFKSTSGS 289

RESULT 4
ALAS3
alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A

C:Species: Aspergillus oryzae
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999

C:Accession: S04549; A33215; A44713
R:Witsell, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
A:Reference number: S04548; MUID:89237897

A:Accession: S04549
A:Molecule type: DNA

A:Residues: 1-499 <WR>
A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922

A:Status: translation not shown
A:Molecule type: mRNA

A:Residues: 1-499 <MT>
A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922

R:Genes, M.J.; Dove, M.J.; Seliy, V.L.
Gene 79, 107-117, 1989

A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each contain
A:Reference number: J50240; MUID:89378767

A:Accession: A44713
A:Molecule type: DNA

A:Residues: 1-499 <GEN>
A:Genetics: AMY3

A>Note: the authors refer to this as isozyme I
R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984

A:Title: Structure and possible catalytic residues of Taka-amylase A.
A:Reference number: A37454; MUID:84212370

C:Contents: annotation; X-ray crystallography, 3.0 angstroms
C:Comment: One atom of calcium per molecule is essential for activity.

C:Genetics: <AMY1>
A:Gene: amy1

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me

F:1-21/Domain: signal sequence #status predicted <Sig>
F:22-499/Product: alpha-amylase 3 #status experimental <MAT>

F:194-321/Domain: alpha-amylase core homology <AMY>
F:51-59, 171-185, 261-304, 461-496/disulfide bonds: #status experimental

F:142, 183, 196, 231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:231, 251, 318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 44; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
 |||||||||
 Db 281 NAFKSTSGS 289

RESULT 5

B48305
 alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
 C:Species: Aspergillus awamori
 C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
 C:Accession: B48305
 R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
 Curr. Genet. 17, 203-212, 1990
 A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Asper
 A:Reference number: A48305; M0ID:90234827
 A:Accession: B48305
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499 <KOR>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 44; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
 |||||||||
 Db 281 NAFKSTSGS 289

RESULT 6

JS0663
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
 C:Species: Aspergillus sp.
 C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
 C:Accession: JS0663
 R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
 Biosci. Biotechnol. Biochem. 56, 174-179, 1992
 A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its expression
 A:Reference number: JS0663; M0ID:92323146
 A:Accession: JS0663
 A:Molecule type: mRNA
 A:Residues: 1-499 <SHI>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <ALP>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 44; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
 |||||||||
 Db 281 NAFKSTSGS 289

RESULT 7

JT0466
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 N:Alternate names: glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
 C:Accession: JT0466
 R:Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
 Agric. Biol. Chem. 53, 593-599, 1989
 A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Asperg
 A:Reference number: JT0466
 A:Accession: JT0466
 A:Molecule type: DNA
 A:Residues: 1-499 <TAD>
 C:Comment: See also PIR:JK0201 and PIR:JS0240.
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics:
 A:Insertions: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 44; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
 |||||||||
 Db 281 NAFKSTSGS 289

RESULT 8

JN0588
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 N:Alternate names: Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
 C:Accession: JN0588
 R:Tsuakoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsuboi, A.; Ueda, S.
 Gene 84, 319-327, 1989
 A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
 A:Reference number: JN0588; M0ID:90128276
 A:Accession: JN0588
 A:Molecule type: mRNA
 A:Residues: 1-499 <TSU>
 C:Comment: The alpha amylases are encoded by multigene family.
 C:Genetics:
 A:Gene: Taa-G1
 A:Insertions: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 44; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9

Db 281 NAFKSTSGS 289

RESULT 9

S17726 H+-transporting ATP synthase (EC 3.6.1.34) beta chain - *Bacillus firmus* (strain CF4)

C:Species: *Bacillus firmus*

C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 19-Jan-2001

C:Accession: S17726

R:iley, D.M.; Krulwich, T.A.

Mol. Gen. Genet. 229, 292-300, 1991

A:Title: Organization and nucleotide sequence of the atp genes encoding the ATP synthase

A:Reference number: S17719; M01D:92017655

A:Accession: S17726

A:Molecule type: DNA

A:Residues: 1-469 <1VE>

A:Cross-references: EMBL:M60117; NID:g2988387; PIDN:AAC08044.1; PID:g142551

C:Genetics:

A:Gene: atpD

C:Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase al

C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; nucleotide binding

F;157-164/Region: nucleotide-binding motif A (P-loop)

F;183-352/Domain: H+-transporting ATP synthase alpha chain homology <ATP>

Query Match 79.5%; Score 35; DB 2; Length 469;

Best Local Similarity 77.8%; Pred. No. 16;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9

Db 28 NAFKSTSGA 36

RESULT 10

S69741

hypothetical protein YDR455c - yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*

C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001

C:Accession: S69741

R:dieterich, F.S.

submitted to the EMBL Data Library, August 1995

A:Description: The sequence of *S. cerevisiae* lambda 3641 and cosmids 9461, 9831, and 941

A:Reference number: S69555

A:Accession: S69741

A:Molecule type: DNA

A:Residues: 1-102 <DIE>

A:Cross-references: EMBL:U33007; NID:g927685; PID:g2194171; GSPDB:GN00004; MIPS:YDR455c

C:Genetics:

A:Gene: MIPS:YDR455c

A:Map position: 4R

C:Superfamily: *Saccharomyces cerevisiae* hypothetical protein YDR455c

Query Match 77.3%; Score 34; DB 2; Length 102;

Best Local Similarity 100.0%; Pred. No. 5.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FKSTSGS 9

Db 15 FKSTSGS 21

RESULT 11

H83888

transmembrane lipoprotein lpIC [imported] - *Bacillus halodurans* (strain C-125)

C:Species: *Bacillus halodurans*

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 14-Dec-2001

C:Accession: H83888

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A:Reference number: A83650; M01D:20512582; PMID:11058132

A:Accession: H83888

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-297 <STO>

A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA005631.1; GSPDB:C

A:Experimental source: strain C-125

C:Genetics:

A:Gene: lpIC

C:Superfamily: maltose transport protein malG

Query Match 75.0%; Score 33; DB 2; Length 297;

Best Local Similarity 75.0%; Pred. No. 26;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSG 8

Db 244 NAFKSTSG 251

RESULT 12

S72270

alpha-amylase (EC 3.2.1.1) precursor - *Cryptococcus* sp. (strain CS2)

C:Species: *Cryptococcus* sp.

A:Variety: strain CS2

C:Date: 23-Apr-1998 #sequence_revision 08-May-1998 #text_change 20-Jun-2000

C:Accession: S72270

R:Ietui, H.; Chino, M.; Kato, M.; Iimura, Y.

Biochem. J. 318, 989-996, 1996

A:Title: Raw-starch-digesting and thermostable alpha-amylase from the yeast *Cryptococ*

A:Reference number: S72270; M01D:96433120

A:Accession: S72270

A:Molecule type: DNA

A:Residues: 1-631 <IEF>

A:Cross-references: EMBL:DB3540; NID:g1595852; PIDN:BAA12010.1; PID:g1595853

A:Experimental source: strain S-2

C:Genetics:

A:Gene: amy-CS2

A:Insertion: 289/2; 326/2

C:Function:

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F;1-20/Domain: signal sequence #status predicted <STG>

F;21-631/Product: alpha-amylase #status predicted <MAT>

F;206-335/Domain: alpha-amylase core homology <AMT>

Query Match 75.0%; Score 33; DB 2; Length 631;

Best Local Similarity 66.7%; Pred. No. 57;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9

Db 295 NAFQSTSGS 303

RESULT 13

A72572

hypothetical protein APE1858 - *Aeropyrum pernix* (strain K1)

C:Species: *Aeropyrum pernix*

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: A72572

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aero*

A:Reference number: A72450; M01D:99310339

A:Accession: A72572

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-100 <RAW>

A:Cross-references: DDBJ:AF000062; NID:g5105244; PIDN:BA80862.1; PID:g5105549
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1858
 C:Superfamily: Aeropyrum pernix hypothetical protein APE1858

Db 176 AFKSTSGS 183

Search completed: June 13, 2002, 08:55:55
 Job time: 375 sec

Query Match 72.7%; Score 32; DB 2; Length 100;
 Best Local Similarity 77.8%; Pred. No. 14;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 ||| ||| |||
 Db 43 NAFKSTSGS 51

RESULT 14
 E70196
 flagellar basal body rod protein homolog flgG - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 20-Aug-1999
 C:Accession: E70196
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Iathigra, R.; White
 son, D.; Peterson, J.; Kervavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 : Bowman, C.; Garland, S.; Fujil, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
 A:Reference number: A70100; MID:98065943
 A:Accession: E70196
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-265 <KDE>
 A:Cross-references: GB:AE001177; GB:AE000783; NID:g2688711; PIDN:AA67132.1; PID:g268872
 A:Experimental source: strain B31
 C:Genetics:
 A:Gene: flgG
 C:Superfamily: rod protein flgG

Query Match 72.7%; Score 32; DB 2; Length 265;
 Best Local Similarity 66.7%; Pred. No. 38;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 ||| ||| |||
 Db 192 NAFKSTSGS 200

RESULT 15
 A39351
 tetragonal surface virulence array protein precursor - Aeromonas salmonicida
 C:Species: Aeromonas salmonicida
 C>Date: 27-Mar-1992 #sequence_revision 10-Apr-1992 #text_change 30-Sep-1993
 C:Accession: A39351
 R:Chu, S.; Cavaignac, S.; Feutrier, J.; Phipps, B.M.; Kostirzyska, M.; Kay, W.W.; Trust,
 J. Biol. Chem. 266, 15258-15265, 1991
 A>Title: Structure of the tetragonal surface virulence array protein and gene of Aeromon
 A:Reference number: A39351; MID:91332048
 A:Accession: A39351
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-502 <CHU>
 A:Cross-references: GB:M64655
 A>Note: the authors translated the codon CCG for residue 98 as Phe

Query Match 72.7%; Score 32; DB 2; Length 502;
 Best Local Similarity 87.5%; Pred. No. 73;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AFKSTSGS 9
 ||| ||| |||

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:53:28 ; Search time 158.68 Seconds
(without alignments)
6.300 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Perfect score: 44

Sequence: 1 NAFKSTSGS 9

Scoring table:

BIOSM62
Gapop 10.0 , Gapept 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	478	15	AA46065
2	44	100.0	478	16	AA46065
3	44	100.0	478	16	AA46065
4	44	100.0	478	16	AA46065
5	44	100.0	478	16	AA46065
6	44	100.0	478	16	AA46065
7	44	100.0	478	16	AA46065
8	44	100.0	478	16	AA46065
9	44	100.0	478	16	AA46065
10	44	100.0	478	16	AA46065
11	44	100.0	478	16	AA46065

12	32	72.7	411	22	ABG24358	Novel human diago
13	32	72.7	1200	22	ABB64803	Drosophila melanog
14	32	72.7	1712	22	ABB44566	Human wound healin
15	31	70.5	367	21	AA495522	Human cyclic nucle
16	31	70.5	473	22	AA66752	Human ATP-dependen
17	31	70.5	473	22	AA678852	Human ATP-dependen
18	31	70.5	474	22	AA67531	Human ATP-dependen
19	31	70.5	490	21	AA195521	Human cyclic nucle
20	31	70.5	500	21	AA43265	Human ORFX ORF3029
21	31	70.5	576	22	AA67533	Amino acid sequenc
22	31	70.5	621	20	AA407029	Breast cancer asso
23	31	70.5	632	17	AA85300	Arabidopsis pathog
24	31	70.5	684	22	AA62678	Human type 11 phos
25	31	70.5	684	22	AA33472	Human cyclic nucle
26	31	70.5	766	22	AA609526	Novel human diago
27	31	70.5	934	22	AA609526	Human type 11 phos
28	31	70.5	1269	17	AA403659	RPP5 downy mildew
29	31	70.5	1391	22	ABB58810	Drosophila melanog
30	30	68.2	116	18	AA420165	H. pylori flagella
31	30	68.2	116	18	AA424596	H. pylori flagella
32	30	68.2	226	21	AA44938	Artichoke sulphate
33	30	68.2	234	18	AA420722	H. pylori flagella
34	30	68.2	387	21	AA49863	Arabidopsis thalia
35	30	68.2	394	22	ABB65646	Drosophila melanog
36	30	68.2	397	22	ABB65647	Drosophila melanog
37	30	68.2	429	22	AA59671	B. napus KAS prote
38	30	68.2	431	22	AA59670	A. thaliana KAS pr
39	30	68.2	441	13	AA42052	Brassica synthase
40	30	68.2	441	15	AA454027	Synthase factor B
41	30	68.2	455	20	AA37419	Chlamydia trachoma
42	30	68.2	499	21	AA41597	Arabidopsis thalia
43	30	68.2	499	21	AA52088	Arabidopsis thalia
44	30	68.2	513	21	AA627321	Arabidopsis thalia
45	30	68.2	534	21	AA631580	Arabidopsis thalia

ALIGNMENTS

RESULT	1
ID	AA46065 standard; protein; 478 AA.
XX	XX
AC	AA46065;
XX	XX
DT	18-JUL-1994 (first entry)
XX	XX
DE	Mutant alpha-amylase.
XX	XX
KW	Methionine substitution; stability; activity; detergent;
KW	dishwashing agents; liguifaction agents.
XX	XX
OS	Aspergillus oryzae.
XX	XX
PN	WO9402597-A.
XX	XX
PD	03-FEB-1994.
XX	XX
PF	06-JUL-1993; 93WO-DK00230.
XX	XX
PR	23-JUL-1992; 92DK-0000946.
PR	16-DEC-1992; 92DK-0001503.
PR	15-MAR-1993; 93DK-0000292.
XX	XX
PA	(NOVO) NOVO-NORDISK AS.
XX	XX
PI	Bisgard-frantzen H, Svendsen A;
XX	XX
DR	WPI; 1994-048855/06.
PT	Mutant alpha-amylase from Bacillus species comprising a
PT	methionine substitution - with improved stability and activity at
PT	low pH, for use in detergents, dishwashing agents and

PT 1liquefaction agents
XX
PS Claim 1; Page 7; 20pp; English.
XX
CC The sequence as that of the *Aspergillus oryzae* alpha amylase, sold
CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC be mutated by substitution of one or more of its methionine residues
CC for any amino acid other than cysteine. The mutant alpha amylase
CC exhibits a better activity level and better stability in the
CC presence of oxidizing agents than previous mutant alpha amylases,
CC and improved thermostability at moderately low pH. The enzyme can
CC be used as an additive for detergents, dishwashing agents and
CC 1liquefaction agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 44; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
Db 260 nafkstsgs 268
|||||

RESULT 2
AAR79025
ID AAR79025 standard; protein; 478 AA.
XX
AC AAR79025;
XX
DT 22-MAR-1996 (first entry)
XX
DE Mature taka-amylase A.
XX
KW Wild type: neopullulanase; B. stearothermophilus; mutant;
KW food industry; modification; hydrophobicity; replacement
KW insertion; deletion.
XX
OS *Aspergillus oryzae*.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
PN JP07177891-A.
XX
PD 18-JUL-1995.
XX
PF 31-OCT-1994; 94JP-0288658.
XX
PR 12-NOV-1993; 93JP-0306096.
XX
PA (EZAK) EZAKI GLICO CO.
PA (NIDE) NEC CORP.
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site - increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
PS Disclosure; Page 10-11; 18pp; English.
XX
CC This sequence represents the mature form of taka-amylase from *A. oryzae*.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the

CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC hydrophobic amino acid from the selected site. The method was used
CC to produce neopullulanases Y377F, S422V and M375L.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 44; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
Db 260 nafkstsgs 268
|||||

RESULT 3
AAR72450
ID AAR72450 standard; protein; 478 AA.
XX
AC AAR72450;
XX
DT 01-DEC-1995 (first entry)
XX
DE *Aspergillus oryzae* alpha amylase (mature protein).
XX
KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KW washing; *Bacillus licheniformis*; *Bacillus amyloliquefaciens*;
KW *Bacillus stearothermophilus*; dyeing; bleaching; scouring; textile;
KW thermostable.
XX
OS *Aspergillus oryzae*.
XX
PN W09510603-A.
XX
PD 20-APR-1995.
XX
PF 05-OCT-1994; 94WO-DK00370.
XX
PR 08-OCT-1993; 93DK-0001133.
PR 02-FEB-1994; 94DK-0000140.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Svendsen A. Thøgersen M;
PI Van der zee P, Bisgardtrantzen H, Borchert T;
XX
DR WPI; 1995-161790/21.
XX
PT New *Bacillus* derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance
XX
PS Disclosure; Page 75-76; 105pp; English.
XX
CC Variant alpha amylase enzymes which have improved washing and/or
CC as detergent additives. The enzymes have one or more amino acid
CC residues added, deleted or substituted. The variants can also be
CC used for textile desizing prior to scouring, bleaching and dyeing.
CC The variants have improved thermostability, acid/alkaline stability;
CC low temperature optimum; pH optimum; higher hydrolysis velocity and
CC improved tolerance to other composition constituents, e.g. oxidation
CC agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 44; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||||


```

FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to this
FT fragment is deleted or replaced with a fragment
FT corresponding to 102-199 of AAM14499, claim 42"
FT
FT Misc-difference 165..177
FT /label= loop 3 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 195-202 of AAM14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 21"
FT
FT Misc-difference 166..173
FT /label= loop 3 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 196-198
FT of AAM14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 23"
FT Misc-difference 181..184
FT /note= "an amino acid fragment corresponding to this
FT region is deleted from the parent sequence of a
FT variant Fungamyl; claim 43"
FT
FT Misc-difference 291..313
FT /label= loop 8 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 322-346 of AAM14498 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 36"
FT
FT Misc-difference 297..313
FT /label= loop 8 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 325-345
FT of AAM14498 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 38"
FT
FT WO9623874-A1.
FT
FT 08-AUG-1996.
FT
FT 05-FEB-1996; 96WO-DK00057.
FT
FT 10-NOV-1995; 95DK-0001256.
FT 03-FEB-1995; 95DK-0000128.
FT 23-OCT-1995; 95DK-0001192.
FT
FT (NOVO ) NOVO-NORDISK AS.
FT
FT Bisgard-frantzen H, Borchert TV, Svendsen A;
FT
FT WPI; 1996-371424/37.
FT
FT Alpha-amylase variants and methods of production - have altered
FT properties such as calcium dependency, substrate binding and
FT stability
FT
FT Disclosure; Page 87-88; 171pp; English.
FT
FT The present sequence is the mature Aspergillus oryzae alpha-amylase (A).
FT Variants of parent termamyl- and fungamyl-like alpha-amylases (and
FT methods of constructing them) are claimed. Examples of variants are
FT featured above. The variants have altered properties such as calcium
FT dependency, substrate binding and stability. Also one or more proline or
FT cysteine residues in the variant is modified or replaced with a
FT non-proline or non-cystein residue such as alanine. The variants can be
FT used for (dis)washing, as detergent additives or for fabric desizing or
FT starch liquetaction. They can also be used for the production of
FT sweeteners and ethanol from starch. See also AAM14498-99.
FT
FT Sequence 478 AA;

```

```

OY Query Match 100.0%; Score 44; DB 17; Length 478;
DB Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
DB 260 nafksts 268

RESULT 6
AAB84206
ID AAB84206 standard; Protein; 498 AA.
XX
AC AAB84206;
XX
DT 06-AUG-2001 (first entry)
XX
DE Amino acid sequence of a fungamyl-like alpha-amylase.
XX
KW Fungamyl-like alpha-amylase; glucoamylase; dextrinisation; maltose;
KW alcohol; starch; dough improver; brewing; starch liquification.
XX
OS Aspergillus oryzae.
XX
PN WO200134784-A1.
XX
PD 17-MAY-2001.
XX
PF 10-NOV-2000; 2000WO-DK00626.
XX
PR 10-NOV-1999; 99DK-0001617.
XX
PA (NOVO ) NOVOZYMS AS.
XX
PI Bisgard-Frantzen H, Svendsen A, Pedersen S;
XX
DR WPI; 2001-367478/38.
XX
DR N-PSDB; AAF90208.
XX
PT New variant of Fungamyl-like alpha-amylase, useful for production of
PT maltose syrups, includes mutations that improve stability against heat
PT and acidic pH -
XX
XX Claim 1; Page 42-45; 49pp; English.
XX
XX The present sequence represents a fungamyl-like alpha-amylase. The
XX specification describes variants of this fungamyl-like alpha-amylase,
XX which have an alteration in one the amino acid regions 98-110, 150-160,
XX 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion
XX or substitution of an amino acid or an insertion of an amino acid
XX downstream of a particular position. The variants retain alpha-amylase
XX activity, and have better heat stability and/or stability at acidic pH,
XX relative to wild-type enzyme. The variants can therefore be used at
XX higher temperatures (more efficient conversion or faster reaction, and
XX have reduced need for cooling and reduced risk of contamination). The
XX variants may also be used in conjunction with other enzymes,
XX particularly glucoamylase during dextrinisation. The variants are
XX used to produce syrups, particularly of high maltose content, or alcohol,
XX from starch; as dough improver for baked goods; in brewing, to increase
XX fermentability of the wort; and for liquetaction of starch.
XX
XX Sequence 498 AA;

```

```

OY Query Match 100.0%; Score 44; DB 22; Length 498;
DB Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
DB 280 nafksts 288

```


KM	Drosophila: developmental biology; cell signalling; insecticide;
KM	pharmaceutical.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PE	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI, 2001-656860/75.
DR	N-PSDB; ABL10155.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
PS	
XX	Disclosure; SEQ ID NO 24948; 21pp + Sequence listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC	sequences (ABP01840-ABL16175) and the encoded proteins
CC	(AB57737-AB812072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
XX	Sequence 690 AA:
QY	
1	NAFKSTGS 9
1	
387	nlfkstss 395
Db	
RESULT 10	
AAB43211	
ID	AAB43211 standard; Protein; 397 AA.
XX	
AC	AAB43211:
DT	
08-FEB-2001	(first entry)
XX	
DE	Human ORFX ORF2975 polypeptide sequence SEQ ID NO:5950.
XX	
KM	Human: open reading frame; ORFX: detection; cytosolic; hepatotropic;
KM	vulnerrary; antipsoiatric; antiparkinsonian; nootropic; neuroprotective;
KM	anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiac;
KM	immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KM	hypotensive; dermatological; immunosuppressive; antiinflammatory;
KM	antiviral; antibacterial; antifungal; antineumatic; antihypoid;
KM	antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KM	neurodegenerative disorder; osteoarthritis; graft vs host disease;
KM	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KM	cholesterol ester storage; systemic lupus erythematosus; infection;
KM	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KM	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;

KM	bone damage; cartilage damage; antinflammatory disease; coagulation;
XW	thrombosis; contraceptive.
OS	Homo sapiens.
XX	MO200058473-A2.
PN	05-OCT-2000.
XX	31-MAR-2000; 2000WO-US08621.
PF	31-MAR-1999; 99US-0127607.
XX	02-APR-1999; 99US-0127636.
PR	03-APR-1999; 99US-0127728.
XX	30-MAR-2000; 2000US-0340763.
PA	(CURA-) CURAGEN CORP.
XX	
PI	Shimkets RA, Leach M;
XX	
DR	WPI: 2000-602362/57.
N-PSDB:	AAC77420.
XX	
PT	Novel nucleic acids and peptides derived from open reading frame X,
PT	useful for treating e.g. cancers, proliferative disorders,
PT	neurodegenerative disorders and cardiovascular disease -
Claim 11;	Page 5133-5134; 5507pp; English.
XX	
XX	AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC	which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC	sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC	antipapillary; antiparkinsonian; nootropic; neuroprotective;
CC	osteoprotic; anticoagulant; antithrombotic; immunosuppressant;
CC	immunostimulant; cardiact; thrombolytic; coagulant; vasotrophic;
CC	antidiabetic; hypotensive; dermatologic; immunosuppressive;
CC	antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC	antihypoid; and antianemic. The sequences can be used for determining
CC	the presence of or predisposition to, or preventing or treating
CC	pathological conditions associated with an ORFX-associated disorder. The
CC	nucleic acids can be used to express ORFX proteins in gene therapy
CC	vectors. The proteins and nucleic acids may be used to treat cancers,
CC	proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC	graft vs host disease, cardiovascular disease, diabetes mellitus,
CC	hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC	erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC	bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC	allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC	nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC	coagulation; to inhibit thrombosis; and as a contraceptive.
XX	
SQ	Sequence 397 AA:
OY	Query Match 72.7%; Score 32; DB 21; Length 397;
DG	Best Local Similarity 75.0%; Pred. No. 1.5e+02;
MATCHES	6; Conservative 1; Mismatches 1; Indels 0; Gaps 0.
OY	1 NAFKSTGS 8
DG	
DG	77 nlfkts 84
RESULT 11	
ID	ABG08887 standard; Protein; 411 AA.
AC	ABG08887;
DT	13-FEB-2002 (first entry)
DE	Novel human diagnostic protein #8878.

KW	Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
PX	WO200175067-A2.
PN	
PD	11-OCT-2001.
PP	
PF	30-MAR-2001: 2001WO-US08631.
PR	31-MAR-2000: 2000US-0540217.
PR	23-AUG-2000: 2000US-0649167.
PA	(HYSB-) HYSBO INC.
PI	Dzmanac RT, Liu C, Tang YT;
DR	WIPI: 2001-639362/73.
DR	N-PDB: AAS73074.
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
PS	Claim 20: SEQ ID No 39246; 103bp; English.
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG30377 represent novel human
CC	diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/Published_pcr_sequences.
SQ	Sequence 411 AA:
OY	1 NAFKSTSGS 9
Db	: :
	381 ntfsssgs 389
RESULT_12	
ABG24358	
ID	ABG24358 standard; Protein; 411 AA.
AC	
XX	ABG24358;
XX	
DT	18-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #24349.
XX	
XX	Human: chromosome mapping; gene mapping; gene therapy; forensic;

KW	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
XX	Homo sapiens.
XX	
XX	WO200175067-A2.
XX	
XX	11-OCT-2001.
XX	
XX	30-MAR-2001; 2001WO-US08631.
XX	
XX	31-MAR-2000; 2000US-0540217.
FR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
XX	Drmnac RT, Liu C, Tang YT;
PI	
XX	WPI: 2001-639362/73.
DR	N-PsDB; AAS88545.
XX	
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
PS	
XX	Claim 20: SEQ ID No 54717; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AB800010-AB830377 represent novel human
CC	diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
XX	
XX	Sequence 411 AA:
SO	
QY	Query Match 72.7%; Score 32; DB 22; Length 411;
DB	Best Local Similarity 66.7%; Pred. NO. 1.6e+02;
Matches	6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY	1 NAFKSTGS 9
	: : :
DB	381 nlfsssgs 389
RESULT	13
ABB64803	
ID	ABB64803 standard; Protein; 1200 AA.
XX	
AC	ABB64803;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 21201.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical.

```
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI: 2001-6556860/75.
XX DR N-PSDB; ABL08906.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Disclosure: SEQ ID NO 21201; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX CC sequences (AB101840-AB16175) and the encoded proteins
XX CC (AB57737-AB572072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at f:ip.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1200 AA;

Query Match 72.7%; Score 32; DB 22; Length 1200;
Best Local Similarity 66.7%; Pred. NO. 5e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
DB 1092 sstestsigs 1100

RESULT 14
AB844566
ID AB844566 standard; Protein; 1712 AA.
XX AC AB844566;
XX DT 25-JAN-2002 (first entry)
XX DE Human wound healing related polypeptide SEQ ID NO 23.
XX KW Human; mouse; vulnery; dermatological; skin disorder; wound healing;
XX KW gene therapy.
XX OS Homo sapiens.
XX PN CA2325226-A1.
XX PD 17-MAY-2001.
XX PF 16-NOV-2000; 2000CA-2325226.
XX PR 17-NOV-1999; 99DE-1055349.
XX PR 17-DEC-1999; 99US-0172511.
XX PR 20-JUN-2000; 2000DE-1030149.
```

```
XX PA (SWIT-) SWITCH BIOTECH AG.
XX PI Regenbogen J, Wolf E, Goppelt A, Werner S, Halle J;
XX DR WPI: 2001-433142/47.
XX PF use of novel polypeptide or its variant or nucleic acid encoding the
XX PT polypeptide for diagnosing and/or preventing and/or treating skin
XX PT disorders and/or treatment in wound healing or for identifying active
XX PT substances -
XX PS Disclosure; Page 158-165; 265pp; English.
XX CC The invention relates to the use of a polypeptide (AB844544-AB844601,
XX CC AB844606-AB844623) or its variant or encoding nucleic acid
XX CC (AB81990-AB81995, AB82016-AB82032) with vulnery and/or
XX CC dermatological activity for the diagnosis, prevention and treatment of
XX CC skin disorders and treatment in wound healing or for the identification
XX CC of pharmacologically active substances. The nucleic acids are useful in
XX CC gene therapy.
XX CC Note: The printed sequence listing for this specification was incomplete,
XX CC terminating part way through SEQ ID NO 106. The remaining data was
XX CC obtained from EPO data for an equivalent patent (EP114862).
XX SQ Sequence 1712 AA;

Query Match 72.7%; Score 32; DB 22; Length 1712;
Best Local Similarity 77.8%; Pred. NO. 7.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
DB 138 nftestsigs 146

RESULT 15
AAY95522
ID AAY95522 standard; Protein; 367 AA.
XX AC AAY95522;
XX DT 10-OCT-2000 (first entry)
XX DE Human cyclic nucleotide phosphodiesterase HSPDE10A2.
XX KW Cyclic nucleotide phosphodiesterase; HSPDE10A2; human; cancer;
XX KW immune disorder; diagnosis; therapy.
XX OS Homo sapiens.
XX PN WO200040733-A1.
XX PD 13-JUL-2000.
XX PF 07-JAN-2000; 2000WO-US00371.
XX PR 07-JAN-1999; 99US-0226741.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Phillips SC, Lanfear J, Fawcett L, Bandman O, Harrow I;
XX DR WPI: 2000-452539/39.
XX DR N-PSDB; AAA49973.
```

PT New human cyclic nucleotide phosphodiesterases, referred as HSPDE10A1
PT and HSPDE10A2, useful for treating and preventing cancer and immune
PT disorders -

XX
PS
XX Claim 1; Page 84; 96pp; English.

CC The present sequence is that of HSPDE10A2, a novel human cyclic
CC nucleotide phosphodiesterase. The amino acid sequence was deduced
CC from an isolated cDNA clone (see AAA49973). HSPDE10A2 is a
CC C-terminal splice variant of HSPDE10A1 (see AAY95521), being
CC identical between amino acid residues 1-338, but different in the
CC C-terminal portion from residue 339 to 367. It shares 40%
CC identity with HSPDE3A1. The invention provides expression vectors,
CC host cells, antibodies, agonists and antagonists, as well as
CC methods for diagnosing, treating or preventing disorders associated
CC with expression of HSPDE10A1, especially cancer and immune
CC disorders.

XX
SQ Sequence 367 AA;

Query Match 70.5%; Score 31; DB 21; Length 367;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||::|||
Db 269 nafqaksgs 277

Search completed: June 13, 2002, 08:53:29
Job time: 279 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:49:15 ; Search time 57.95 Seconds
(without alignments) 209,904 Million cell updates/sec

Title: US-09-710-339-2
Perfect score: 2684
Sequence: 1 MWAMWSLFLYGLVAPALAA.....LPRVLYPTERLAKSGKICSSS 498

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCBUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2572	95.8	478	1	US-08-720-899-7
2	2572	95.8	478	1	US-08-459-610-7
3	2572	95.8	478	2	US-08-343-804-7
4	2572	95.8	478	2	US-08-600-908A-10
5	2572	95.8	478	3	US-08-683-838A-10
6	2572	95.8	478	4	US-09-182-859-7
7	2503.5	93.3	478	2	US-08-339-715A-2
8	1392	51.9	468	1	US-08-470-702-6
9	1392	51.9	468	1	US-08-467-831-6
10	1391	51.8	468	1	US-08-204-656B-4
11	1391	51.8	468	1	US-08-470-702-7
12	1391	51.8	468	1	US-08-467-831-7
13	1388	51.7	468	1	US-08-204-656B-6
14	1388	51.7	468	1	US-08-470-702-8
15	1387	51.7	468	1	US-08-467-831-8
16	1387	51.7	468	1	US-08-204-656B-2
17	1387	51.7	468	1	US-08-204-656B-8
18	1387	51.7	468	1	US-08-470-702-9
19	1387	51.7	468	1	US-08-467-831-9
20	450	16.8	719	4	US-09-386-607-2
21	446	16.6	686	3	US-08-947-965-73
22	444	16.5	685	3	US-08-947-965-72
23	428.5	16.0	655	1	US-08-469-202-28
24	428.5	16.0	655	2	US-08-484-434C-35
25	427.5	15.9	655	1	US-08-469-202-27
26	426.5	15.9	655	2	US-08-484-434C-34
27	426.5	15.9	683	3	US-08-947-965-2

28	421.5	15.7	624	3	US-08-947-965-78	Sequence 78, Appl
29	418	15.6	676	3	US-08-947-965-71	Sequence 71, Appl
30	413	15.4	686	3	US-08-947-965-70	Sequence 70, Appl
31	413	15.4	687	3	US-08-947-965-75	Sequence 75, Appl
32	404.5	15.1	833	4	US-09-514-302-3	Sequence 74, Appl
33	404.5	15.1	1938	4	US-09-514-302-2	Sequence 76, Appl
34	398	14.8	685	3	US-08-947-965-74	Sequence 77, Appl
35	396	14.8	675	3	US-08-947-965-76	Sequence 5, Appl
36	385.5	14.4	680	3	US-08-947-965-77	Sequence 1, Appl
37	370	13.8	588	2	US-08-339-715A-5	Sequence 4, Appl
38	368	13.7	588	2	US-08-339-715A-1	Sequence 10, Appl
39	367.5	13.7	725	2	US-08-816-105A-1	Sequence 10, Appl
40	367	13.7	588	2	US-08-339-715A-4	Sequence 10, Appl
41	367	13.7	687	1	US-08-470-702-10	Sequence 10, Appl
42	367	13.7	687	1	US-08-467-831-10	Sequence 10, Appl
43	366	13.6	687	1	US-08-204-656B-10	Sequence 3, Appl
44	364	13.6	588	2	US-08-339-715A-3	Sequence 2, Appl
45	259.5	9.7	686	3	US-09-306-922-2	

ALIGNMENTS

RESULT 1
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchart, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720, 899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343, 804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmeyer Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054, 214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7
Query Match 95.8%; Score 2572; DB 1; Length 478;
Best local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;


```

;
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmy Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-343-804-7

Query Match          95.8%; Score 2572; DB 2; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADMRSSQSYFLTLDFRARTDSTTATCNTADOKYCGGTWGGIIDLKDIYIGMGFTAI 80
DB 1 ATPADMRSSQSYFLTLDFRARTDSTTATCNTADOKYCGGTWGGIIDLKDIYIGMGFTAI 60
QY 81 WITPPTAQLPQTTAYGDAYHGYWQODIYSLNENYGTADDLKALSSALHERGMTLWVDVYA 140
DB 61 WITPPTAQLPQTTAYGDAYHGYWQODIYSLNENYGTADDLKALSSALHERGMTLWVDVYA 120
QY 141 NHMGYDAGSSVDYSVFPPSSODYFHPFCFIQNEYEDQTOVEECWGLDNTVSLPDLDTTK 200
DB 121 NHMGYDAGSSVDYSVFPPSSODYFHPFCFIQNEYEDQTOVEECWGLDNTVSLPDLDTTK 180
QY 201 DVKKNEMDWGSLVSNYSIDGLRIDTVKHYQKDFWPGYNKKAAGVYCIQEVLDGDPAYTC 260
DB 181 DVKKNEMDWGSLVSNYSIDGLRIDTVKHYQKDFWPGYNKKAAGVYCIQEVLDGDPAYTC 240
QY 261 PYQNMDCVLNPIYPLPLNAFKSTSGMDLYNMINTVASCDCPSDSTLGLTFEVENHNDNR 320
DB 241 PYQNMDCVLNPIYPLPLNAFKSTSGMDLYNMINTVASCDCPSDSTLGLTFEVENHNDNR 300
QY 321 FASYNDIALAKNVAAFIILNDGIPITAYAGEQHYAGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYNDIALAKNVAAFIILNDGIPITAYAGEQHYAGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNVAISKDGFVYTKNMPYIKDDITIAMKGTGDSQIYIILSNKASGDSYT 440
DB 361 LIASANAIRNVAISKDGFVYTKNMPYIKDDITIAMKGTGDSQIYIILSNKASGDSYT 420
QY 441 LSLSGAGYTAGOQLFEVIGCTTIVGSDGNVPVMAGGLPRVLYPTKLAGSKICSS 498
DB 421 LSLSGAGYTAGOQLFEVIGCTTIVGSDGNVPVMAGGLPRVLYPTKLAGSKICSS 478

RESULT 4
US-08-600-908A-10
; Sequence 10, Application US/08600908A
; Patent No. 5989169
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: "-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5989169 No. 5989169disk of No. 5989169th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,908A
; FILING DATE: 13-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-08-600-908A-10

Query Match          95.8%; Score 2572; DB 2; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADMRSSQSYFLTLDFRARTDSTTATCNTADOKYCGGTWGGIIDLKDIYIGMGFTAI 80
DB 1 ATPADMRSSQSYFLTLDFRARTDSTTATCNTADOKYCGGTWGGIIDLKDIYIGMGFTAI 60
QY 81 WITPPTAQLPQTTAYGDAYHGYWQODIYSLNENYGTADDLKALSSALHERGMTLWVDVYA 140
DB 61 WITPPTAQLPQTTAYGDAYHGYWQODIYSLNENYGTADDLKALSSALHERGMTLWVDVYA 120
QY 141 NHMGYDAGSSVDYSVFPPSSODYFHPFCFIQNEYEDQTOVEECWGLDNTVSLPDLDTTK 200
DB 121 NHMGYDAGSSVDYSVFPPSSODYFHPFCFIQNEYEDQTOVEECWGLDNTVSLPDLDTTK 180
QY 201 DVKKNEMDWGSLVSNYSIDGLRIDTVKHYQKDFWPGYNKKAAGVYCIQEVLDGDPAYTC 260
DB 181 DVKKNEMDWGSLVSNYSIDGLRIDTVKHYQKDFWPGYNKKAAGVYCIQEVLDGDPAYTC 240
QY 261 PYQNMDCVLNPIYPLPLNAFKSTSGMDLYNMINTVASCDCPSDSTLGLTFEVENHNDNR 320
DB 241 PYQNMDCVLNPIYPLPLNAFKSTSGMDLYNMINTVASCDCPSDSTLGLTFEVENHNDNR 300
QY 321 FASYNDIALAKNVAAFIILNDGIPITAYAGEQHYAGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYNDIALAKNVAAFIILNDGIPITAYAGEQHYAGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNVAISKDGFVYTKNMPYIKDDITIAMKGTGDSQIYIILSNKASGDSYT 440
DB 361 LIASANAIRNVAISKDGFVYTKNMPYIKDDITIAMKGTGDSQIYIILSNKASGDSYT 420
QY 441 LSLSGAGYTAGOQLFEVIGCTTIVGSDGNVPVMAGGLPRVLYPTKLAGSKICSS 498
DB 421 LSLSGAGYTAGOQLFEVIGCTTIVGSDGNVPVMAGGLPRVLYPTKLAGSKICSS 478

RESULT 5
US-08-683-838A-10
; Sequence 10, Application US/08683838A
; Patent No. 6022724
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: "-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6022724 No. 6022724disk of No. 6022724th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
```

```
City: New York
State: New York
Country: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394,204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-683-838A-10
```

```
Query Match          95.8%; Score 2572; DB 3; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 21 ATPADMRSGSIYFLTDRFARTDGTATCNTADOKYCGGTWOGIIDKLDYIOGMGFTAI 80
DB 1 ATPADMRSGSIYFLTDRFARTDGTATCNTADOKYCGGTWOGIIDKLDYIOGMGFTAI 60
QY 81 WTPVTAOLPOTTAAGDAHGYWODIYSLNENGTADDLKALSALHERGMTLVNVA 140
DB 61 WTPVTAOLPOTTAAGDAHGYWODIYSLNENGTADDLKALSALHERGMTLVNVA 120
QY 141 NHGYDGASSVDYVFEKPFSSODYFHFPCFIQNYEDQYVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHGYDGASSVDYVFEKPFSSODYFHFPCFIQNYEDQYVEDCWLGDNTVSLPDLDTTK 180
QY 201 DVKKNEMTWGSLVSNYSIDGLRIDYKHKQKDFWPGYNKAAGYVCIGEVLDGDPAYTC 260
DB 181 DVKKNEMTWGSLVSNYSIDGLRIDYKHKQKDFWPGYNKAAGYVCIGEVLDGDPAYTC 240
QY 261 PYQNVMDGYLVNPIYYPPLNFAKSTSGSMDDLNNMINIVKSDCPDSTLLGTFVENHDNPR 320
DB 241 PYQNVMDGYLVNPIYYPPLNFAKSTSGSMDDLNNMINIVKSDCPDSTLLGTFVENHDNPR 300
QY 321 FASYTNDIALAKNVAFIILNDGIPITYAGQBOHYAGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAFIILNDGIPITYAGQBOHYAGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNYAISKDTGVTYKKNMPIYKDDTTIAMRGKTGDSQIVTILSKKSGSGSYT 440
DB 361 LIASANAIRNYAISKDTGVTYKKNMPIYKDDTTIAMRGKTGDSQIVTILSKKSGSGSYT 420
QY 441 LSLSGAGYTAGOOLTEVIGCTTVVSGDGNVPVPAAGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGOOLTEVIGCTTVVSGDGNVPVPAAGLPRVLYPTEKLAGSKICSSS 478
```

```
RESULT 6
US-09-182-859-7
; Sequence 7, Application US/09182859
```

```
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchardt, Jorden
APPLICANT: Bisgaard-Frantzen, Henrik
TITLE OF INVENTION: Alpha-Amylase Mutants
FILE REFERENCE: 4796,204-US
CURRENT APPLICATION NUMBER: US/09/182,859
CURRENT FILING DATE: 1998-10-29
EARLIER APPLICATION NUMBER: 0515/96
EARLIER FILING DATE: 1996-04-30
EARLIER APPLICATION NUMBER: 0712/96
EARLIER FILING DATE: 1996-06-28
EARLIER APPLICATION NUMBER: 0775/96
EARLIER FILING DATE: 1996-07-11
EARLIER APPLICATION NUMBER: 1263/96
EARLIER FILING DATE: 1996-11-08
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 478
TYPE: PRP
ORGANISM: Bacillus licheniformis
US-09-182-859-7
```

```
Query Match          95.8%; Score 2572; DB 4; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 21 ATPADMRSGSIYFLTDRFARTDGTATCNTADOKYCGGTWOGIIDKLDYIOGMGFTAI 80
DB 1 ATPADMRSGSIYFLTDRFARTDGTATCNTADOKYCGGTWOGIIDKLDYIOGMGFTAI 60
QY 81 WTPVTAOLPOTTAAGDAHGYWODIYSLNENGTADDLKALSALHERGMTLVNVA 140
DB 61 WTPVTAOLPOTTAAGDAHGYWODIYSLNENGTADDLKALSALHERGMTLVNVA 120
QY 141 NHGYDGASSVDYVFEKPFSSODYFHFPCFIQNYEDQYVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHGYDGASSVDYVFEKPFSSODYFHFPCFIQNYEDQYVEDCWLGDNTVSLPDLDTTK 180
QY 201 DVKKNEMTWGSLVSNYSIDGLRIDYKHKQKDFWPGYNKAAGYVCIGEVLDGDPAYTC 260
DB 181 DVKKNEMTWGSLVSNYSIDGLRIDYKHKQKDFWPGYNKAAGYVCIGEVLDGDPAYTC 240
QY 261 PYQNVMDGYLVNPIYYPPLNFAKSTSGSMDDLNNMINIVKSDCPDSTLLGTFVENHDNPR 320
DB 241 PYQNVMDGYLVNPIYYPPLNFAKSTSGSMDDLNNMINIVKSDCPDSTLLGTFVENHDNPR 300
QY 321 FASYTNDIALAKNVAFIILNDGIPITYAGQBOHYAGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAFIILNDGIPITYAGQBOHYAGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNYAISKDTGVTYKKNMPIYKDDTTIAMRGKTGDSQIVTILSKKSGSGSYT 440
DB 361 LIASANAIRNYAISKDTGVTYKKNMPIYKDDTTIAMRGKTGDSQIVTILSKKSGSGSYT 420
QY 441 LSLSGAGYTAGOOLTEVIGCTTVVSGDGNVPVPAAGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGOOLTEVIGCTTVVSGDGNVPVPAAGLPRVLYPTEKLAGSKICSSS 478
```

```
RESULT 7
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
GENERAL INFORMATION:
APPLICANT: Kaneko, Hiroki
APPLICANT: Takada, Toshikazu
APPLICANT: Shimada, Jiro
APPLICANT: Kuriki, Takashi
APPLICANT: Yanase, Michiyo
```



```

? APPLICANT: Takata, Hiroki
? TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
? TITLE OF INVENTION: NEOPULULANASE
? NUMBER OF SEQUENCES: 7
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: POPHAM, HAIR, SCHNORRICH & KAUFMAN, LTD.
? STREET: Metropolitan Square Building, Suite 800, 1450
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20005
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/339,715A
? FILING DATE: 14-NOV-1994
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 306096/1993
? FILING DATE: 12-NOV-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Meiser, Allen S.
? REGISTRATION NUMBER: 27,215
? REFERENCE/DOCKET NUMBER: 18335,009
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-824-8000
? TELEFAX: 202-824-8199
? TELEX: 248516
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 478 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-339-715A-2

```

```

Query Match          93.3%; Score 2503.5; DB 2; Length 478;
Best Local Similarity 97.9%; Pred. No. 1.5e-214;
Matches 468; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 21 ATPADMRSGSITFLTDRFARTDGSITATCNTADOKYCGGTWOGIIDKLDYIYQGMGFTAI 80
DB 1 ATPADMRSGSITFLTDRFARTDGSITATCNTADOKYCGGTWOGIIDKLDYIYQGMGFTAI 60

QY 81 WITPVTAQLPOTTAAGDAVHYGYWODIYSLNENYGTADDLKAISALHERGMATLVADYVA 140
DB 61 WITPVTAQLPOTTAAGDAVHYGYWODIYSLNENYGTADDLKAISALHERGMATLVADYVA 120

QY 141 NHMGYDAGSSVDYVYFKPFSSQDYFHPFCFIONYEDQTOVEDCWLGDNTVSLPLDITTK 200
DB 121 NHMGYDAGSSVDYVYFKPFSSQDYFHPFCFIONYEDQTOVEDCWLGDNTVSLPLDITTK 180

QY 201 DYYKNEWDWGSISYNSISGLRIDTYKHYQKDFWPGYKNAAGYCCIGEVLDGDPAYTC 260
DB 181 DYYKNEWDWGSISYNSISGLRIDTYKHYQKDFWPGYKNAAGYCCIGEVLDGDPAYTC 240

QY 261 PYQNVMDGVLANPIYYPILNAFKSTSGSMDLYNMINTVKSDCPSDSTLGFVEVHNHDPK 320
DB 241 PYQNVMDGVLANPIYYPILNAFKSTSGSMDLYNMINTVKSDCPSDSTLGFVEVHNHDPK 300

QY 321 PASTYNDIALAKNVAFTIILNDGPIIYAGQEOHYAGGNDPANREATWLSGYPTDSELYK 380
DB 301 PASTYNDIALAKNVAFTIILNDGPIIYAGQEOHYAGGNDPANREATWLSGYPTDSELYK 360

QY 381 LIASANAIRNVAISKDQGFVYTKNMPYIKDDTTIAMRGCTGSGQIVTILSNKKGASGDSYT 440
DB 361 LIASANAIRNVAISKDQGFVYTKNMPYIKDDTTIAMRGCTGSGQIVTILSNKKGASGDSYT 440

```

```

DB 361 LIASANAIRNVAISKDQGFVYTKN-PIKDDTTIAMRGCTGSGQIVTILSNKKGASGDSYT 419
QY 441 LISAGAGTAQOQLEVIYIGCTTVVSGSNVPVMAGGLPRVLYPTEKLAGSKICSS 498
DB 420 LISGASYTAQOQLEVIYIGCTTVVSGSNVPVMAGGLPRVLYPTEKLAGSKICSDS 477

```

```

RESULT 8
US-08-470-702-6
? Sequence 6, Application US/08470702
? Patent No. 5631149
? GENERAL INFORMATION:
? APPLICANT: MATSUI, IKUO
? APPLICANT: ISHIKAWA, KAZUHIKO
? APPLICANT: MIYAIRI, SACHIO
? APPLICANT: HONDA, KOICHI
? TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
? TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
? STREET: 8110 Gatehouse Road, Suite 500 East
? CITY: Falls Church
? STATE: Virginia
? COUNTRY: U.S.A.
? ZIP: 22042
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/470,702
? FILING DATE: 06-JUN-1995
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/204,656
? FILING DATE: 02-MAR-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: WEINER, MARC S.
? REGISTRATION NUMBER: 32,181
? REFERENCE/DOCKET NUMBER: 234-252P
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 205-8000
? TELEFAX: (703) 205-8050
? TELEX: 248345
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 468 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? US-08-470-702-6

```

```

Query Match          51.9%; Score 1392; DB 1; Length 468;
Best Local Similarity 54.4%; Pred. No. 1.2e-115;
Matches 259; Conservative 81; Mismatches 126; Indels 10; Gaps 5;

QY 22 TPAD-WMSQSIYFLTDRFARTDGSITATCNTADOKYCGGTWOGIIDKLDYIYQGMGFTAI 80
DB 2 TPAD-WMSQSIYFLTDRFARTDGSITATCNTADOKYCGGTWOGIIDKLDYIYQGMGFTAI 61

QY 81 WITPVTAQLPOTTAAGDAVHYGYWODIYSLNENYGTADDLKAISALHERGMATLVADYVA 140
DB 62 WISPVYENITDNATGAYHYGFMKKNITKINENFGTAADLKSLAQELHPRDMLVADYV 121

QY 141 NHMGYDAGSSVDYVYFKPFSSQDYFHPFCFIONYEDQTOVEDCWLGDNTVSLPLDITTK 200
DB 121 NHMGYDAGSSVDYVYFKPFSSQDYFHPFCFIONYEDQTOVEDCWLGDNTVSLPLDITTK 200

```

Db 122 NHGSDGSGSDIDSEYTPENDOKTFHNHCLISNYDDQAQVOSCWEGSDSSVALPDLRTED 181

QY 201 DVYKNEYMDVWGSILVNSYIDGLRIDTVKHYOKDFWPGYNKAGVYCIAGEVLDDPAYTC 260

Db 182 SDVASVFNMSWKDFVGNYSIDGLRIDSADKHYDGFDFPVASAGVYSGVEFGODDPAYTC 241

QY 261 PYQNVMDGVLNPIYPLNAFKSTSGSMDLYNMINTVKSDECPDSTLLGTFVENHNDNR 320

Db 242 PYQNVYIPGVSNPIYPTTREFKTTDSSSELQMISSVASSCDPTLLTNFVENHNDNR 301

QY 321 FASYNDIALAKVNAFFIINDGIPITVAGOEQHYAGGNDPANREATWLSGYPDSELYK 380

Db 302 FASMTSDSLISNALAFVLGDGIPVITYGQEGLSGKSDPNREALMSGYNKESDYK 361

QY 381 LIASANAIRNVAISKDTGFTYTKMNPYIKDPTTIAMRKGTGDSQIYTLISNKGASGDSYT 440

Db 362 LIAKANAARNAVYODSSVATSSQLSVIFSNHDVIAIKRGS----VVSFNNLGSSGSS-D 416

QY 441 LSLSGAGTYAGQOLTEYIGCTTVYVGSOGNVPVPMAGLPVIVLPTKLAGSKICS 496

Db 417 VTISNTGYSSEDEVLVLTCTVSGSSD--LQVSIQGGQPIFVPAK--YASDICS 468

RESULT 9

US-08-467-831-6

/ Sequence 6, Application US/08467831

/ Patent No. 5635378

/ GENERAL INFORMATION:

/ APPLICANT: MATSUI, IKUO

/ APPLICANT: ISHIKAWA, KAZUHIKO

/ APPLICANT: MIYAIRI, SACHIO

/ APPLICANT: HONDA, KOICHI

/ TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,

/ TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING

/ NUMBER OF SEQUENCES: 17

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH

/ STREET: 8110 Gatehouse Road, Suite 500 East

/ CITY: Falls Church

/ STATE: Virginia

/ COUNTRY: U.S.A.

/ ZIP: 22042

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/467,831

/ FILING DATE: 06-JUN-1995

/ CLASSIFICATION: 435

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 08/204,656

/ FILING DATE: 02-MAR-1994

/ ATTORNEY/AGENT INFORMATION:

/ NAME: WEINER, MARC S.

/ REGISTRATION NUMBER: 32,181

/ REFERENCE/DOCKET NUMBER: 234-252P

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (703) 205-8000

/ TELEFAX: (703) 205-8050

/ TELEX: 248345

/ INFORMATION FOR SEQ ID NO: 6:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 468 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

/ HYPOTHEICAL: NO

/ ANTI-SENSE: NO

/ US-08-467-831-6

Query Match 51.9%; Score 1392; DB 1; Length 468;

Best Local Similarity 54.4%; Pred. No. 1,2e-113;

Matches 259; Conservative 81; Mismatches 126; Indels 10; Gaps 5;

QY 22 TPAD-WRSOSIFLLDRFARPTGCTTATCMTADOKYCGGTWOGIIDKIDYIQMGFTAI 80

Db 2 TNADKMRSSOTIYVIDRRAKRDGTSASCNTEDRLYCGSGOIGIKKIDYIKDMGFTAI 61

QY 81 WTPVTAOLPOTTAAYDAYHGYWOODIYSLNENYGTADDLKALSSALHERGYLAWDYVA 140

Db 62 WISPVVENIUDNTAAYGAAHGFEMKNIYKINENFGTADDLKSLADELHSDMLAWDIYT 121

QY 141 NHGTYGAGSVDYVYFKEPPSSODYFHPFCFTONTYEDQVOVEDCMTGDNVLSLPDLDTK 200

Db 122 NHGSDGSGSDIDSEYTPENDOKTFHNHCLISNYDDQAQVOSCWEGSDSSVALPDLRTED 181

QY 201 DVYKNEYMDVWGSILVNSYIDGLRIDTVKHYOKDFWPGYNKAGVYCIAGEVLDDPAYTC 260

Db 182 SDVASVFNMSWKDFVGNYSIDGLRIDSADKHYDGFDFPVASAGVYSGVEFGODPAYTC 241

QY 261 PYQNVMDGVLNPIYPLNAFKSTSGSMDLYNMINTVKSDECPDSTLLGTFVENHNDNR 320

Db 242 PYQNVYIPGVSNPIYPTTREFKTTDSSSELQMISSVASSCDPTLLTNFVENHNDNR 301

QY 321 FASYNDIALAKVNAFFIINDGIPITVAGOEQHYAGGNDPANREATWLSGYPDSELYK 380

Db 302 FASMTSDSLISNALAFVLGDGIPVITYGQEGLSGKSDPNREALMSGYNKESDYK 361

QY 381 LIASANAIRNVAISKDTGFTYTKMNPYIKDPTTIAMRKGTGDSQIYTLISNKGASGDSYT 440

Db 362 LIAKANAARNAVYODSSVATSSQLSVIFSNHDVIAIKRGS----VVSFNNLGSSGSS-D 416

QY 441 LSLSGAGTYAGQOLTEYIGCTTVYVGSOGNVPVPMAGLPVIVLPTKLAGSKICS 496

Db 417 VTISNTGYSSEDEVLVLTCTVSGSSD--LQVSIQGGQPIFVPAK--YASDICS 468

RESULT 10

US-08-204-656B-4

/ Sequence 4, Application US/08204656B

/ Patent No. 5538882

/ GENERAL INFORMATION:

/ APPLICANT: Matsui, Ikuo

/ APPLICANT: Ishikawa, Kazuhiko

/ APPLICANT: Miyairi, Sachio

/ APPLICANT: Honda, Koichi

/ TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,

/ TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing

/ NUMBER OF SEQUENCES: 18

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP

/ STREET: 8110 Gatehouse Road, Suite 500 East

/ CITY: Falls Church

/ STATE: Virginia

/ COUNTRY: U.S.A.

/ ZIP: 22042

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/204,656B

/ FILING DATE: 02-MAR-1994

/ CLASSIFICATION: 435

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Weiner, Marc S.

/ REGISTRATION NUMBER: 32,181

/ REFERENCE/DOCKET NUMBER: 234-252P

/ TELECOMMUNICATION INFORMATION:

```

? TELEPHONE: (703) 205-8000
? TELEFAX: (703) 205-8050
? TELEX: 246345
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 468 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? OS-08-204-656B-4

```

Query Match	51.8%	Score 1391	DB 1	Length 468
Best Local Similarity	54.4%	Pred. No. 1.5e+15		
Matches 259	Conservative 81	Mismatches 126	Indels 10	Gaps 5

QY	22	TPAD-WRSQSIYELLJDRFARTDOSTATCNCTADOKKCGSGWCIIDKDIYIOGMETAI	80
Db	2	TNAKMSQSIOYQIVDRFARTDSDTASACNTERLACGSGSFQI IKKLYIKMGFTAI	61
QY	81	WIFPVTAQLPOTAYDGAHYGWOODIYLSINENGTADDLKALSALHEGMYLMDVVA	140
Db	62	WISPEVERIPDNTAYGAYAHGWMWKKNIKYINENGTADDLKLSLAQELHDDMLMDVIY	121
QY	141	NHMGTGAGSSVDYSVKRPSSODYFHPFCFIQNYEDQTOVEDCWLGDNTVSLPDLDTTK	200
Db	122	NHYSDSGSGSIDSEYTFEPDOKYFFNNYCLISNYDDOAOVOSCEGSDSSVALPDLTRED	181
QY	201	DYVKNENYDWGYSLSVNSYIDGLRIDPHYKHOKEFMYKKAGAYCIGEVILGDPAVYC	266
Db	182	SDVASVNSMWKDPVGNYSIDGLRIDSAKHVDOGFPEFDPASASVYSVGEYFOGDPAYTC	241
QY	261	PYQVWMDGVLANYPYIYLLNNAFKSTSGSMDLYNMINTVKSDCPDSTLLGTFEYENHDNPR	320
Db	242	PYQYIYIGVSNYPYIYPTFFRTKTTDSSSELQMSISVASSCSDPILLTFNFEYHNDR	301
QY	321	FASYTNDIALAKNVAAPILIINDGPIPIYAGOEBOHYAGCNPAPREATWISGYPITDSLYK	380
Db	302	FASWTSSQSLSNALIAVILLDGIPVIYIOQEOGLSKSPNNRREALWLSGYKNESDYK	361
QY	381	LIASANAIIRNVAISKDTEGYVYTKNMPIYKDDTTIAMKRGDGSIOYILISNKAASGDSYT	440
Db	362	LIAMKANARAAVAAYQDSSVYATSQLSVIFFSNDHVIATKRGSG---VVSVFNNLSSGSS-D	416
QY	441	LSISGAGYITAGOOOLEYIGCTIYVSGDNVPPMAGLEPRVLYPIETKLASGYICS	496
Db	417	VTISNTYSSGDELEVILTSIVSGSSD--LYVSIQGGQJQIVPAK--YASDICS	468

RESULT 11
 US-08-470-702-7
 ; Sequence 7, Application US/08470702
 ; Patent No. 5631149
 ; GENERAL INFORMATION:
 ; APPLICANT: MATSUI, IKUO
 ; APPLICANT: ISHIKAWA, KAZUHIKO
 ; APPLICANT: MIYAIRI, SACHIO
 ; APPLICANT: HONDA, KOICHI
 ; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
 ; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
 ; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 ; STREET: 8110 Gatehouse Road, Suite 500 East
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22042
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
ADDITIONAL COMMENTS: 8/26/98 14:23 FAX

```
; INFORMATION FOR SEQ ID NO: 7:
;
; SEQUENCE CHARACTERISTICS:
```

US-08-470-702-7

Query Match	51.8%;	Score 1391;	DB 1;	Length 468;
Best Local Similarity	54.4%;	Pred. No. 1.5e-115;		
Matches 259;	Conservative 81;	Mismatches 126;	Indels 10;	Gaps 5;

QY	22	TRPD -WRSOIYFLLTDRFARIDGSTRATATCTADQKICGSGTGOIIDLKDYIOGMEFTAI	80
QY	22	TRPD -WRSOIYFLLTDRFARIDGSTRATATCTADQKICGSGTGOIIDLKDYIOGMEFTAI	80
Db	2	TRNDKRSOSITQIYTDPRFARDGDSASCSTEDRLCSGSGFGGIIKKDYIDMGFTAI	61
QY	81	WTPPYAQLPQTTAAVDGAHVHGOODIYSLNENGTADDLKALSALHBERGMIAMDVA	14
QY	62	WISPVENIPDNTAAGVAVHGMWKKNIYKINENGTADDLKLSLAQELHBRDMLAMDYIT	12
Db	141	NHMGYDGAAGSSVDYVEKPFSSODYFHPEFCTIOWEDQTOVEDCWLGDNTVSPLDITTK	20
QY	122	NHSGSGSGSDSIDYSEYPRPFNDOKYFHNHYCLISNYDDQAOVOSCWEGSDSVALPDLRTED	18
Db	201	DVYKNEMYMWSILSYNYSIDGLRIDYKHQKCFMGVYKAGVYCISEYIDGDPATYC	26
QY	182	SDVASFENSWADEFGWNSIDGLKIDSXKHVDSQFPDFVSASGVYVSGEYFQGDPAITC	24
Db	261	PIQNWQCVLNTPIYPLLNAFKSTSGSMDDLYNMIINTVAKSDCPDSTLLGTVEVHNDR	32
QY	242	PIQNYIPIGVSNTPIYPTTRFFKTTDSSSELQDMSIVASSCSDDPTLLTNFEVHNDR	30
Db	321	PASYNIDIALAKNVAFFIINDGIPITIIYAGOEQYAGNNDPAREKNTWLSGYPTDSELYK	38
QY	302	PASMTSDOSLISNAIAFYLLBDGIPVITYGEOGSLKSGSDPNRREKMLWLSGYNKESDYK	36
Db	381	LIASANAIRNAISKDTGFEVTYKMWPIYKDDTIYAMKSGPDGSOIYITLSNKGASGDSYT	44
QY	362	LIAKNAANARNAAYVODSSYATSSQLSYIESNDHVIATRKGS---VYSVFNLLGSSGSS-D	41
Db	441	LISLQAGYTAQOOLEVYIGCTTYVYSGDGNVPPMAGCLPRVLYPIPEKLAGSKICS	496
QY	417	VITSNYTGSSGDEIVLEVTLCSTGSSD--LQVSIQGGQIIVPAPK--LYADSDS	468

RESULT 12
US-08-467-831-7
Sequence 7, Application US/084678312
Patent No. 5635378
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIKI, SACHIO

```

? APPLICANT: HONDA, KOICHI
? TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
? TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
? TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
? STREET: 8110 Gatehouse Road, Suite 500 East
? CITY: Falls Church
? STATE: Virginia
? COUNTRY: U.S.A.
? ZIP: 22042
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/467,831
? FILING DATE: 06-JUN-1995
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/204,656
? FILING DATE: 02-MAR-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: WEINER, MARC S.
? REGISTRATION NUMBER: 32,181
? REFERENCE/DOCKET NUMBER: 234-252P
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 205-8000
? TELEFAX: (703) 205-8050
? TELEX: 248345
? INFORMATION FOR SEQ. ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 468 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHEICAL: NO
? ANTI-SENSE: NO
? US-08-467-831-7

```

```

Query Match 51.8%; Score 1391; DB 1; Length 468;
Best Local Similarity 54.4%; Pred. No. 1.5e-115;
Matches 259; Conservative 81; Mismatches 126; Indels 10; Gaps 5;

QY 22 TPAD-WRSQSIYFLTDFRARTDGTATCMTADOKYCGGTGOGILDKDYIQMGFTAI 80
DB 2 TNADKMRQSIYQIVTDFARTDGTASCTEDRIKYGSGFQGIKKLDYIKMGFTAI 61
QY 81 WTPVTAQLPOTAYAGYHGYWODIYSLNENYGTADDLKALSLAHERGMYLMDVVA 140
DB 62 WISPVENIPNTAGYAGYHGMKMKNIYKINENFTADDLKSLAQLHRLMDLMDVIVT 121
QY 141 NHMGYDAGSSVDYSEVFEKPFSSQDYFHPFCFIONYEDQTOVEDCWLGDNTVSLPDLDTK 200
DB 122 NHYGSDDSGSDIDYSEVTEPFNDQKYFHNKCLISNDDQAOVQSCWEGSSVALPDLFRED 181
QY 201 DVKRKNEDYDWGSLVNSYSDGLRIDTVKHYQKDFWPGYNKAAGYVTCGEVLDGDPAYTC 260
DB 182 SDVASVFNKWDYFNKSIDGLRIDSAKHYDQGFPEFVSASGVYSGVEFGQDPAYTC 241
QY 261 PYQNTMDGVNPIYIYPLLNFKSTSGSMDLYNNINTVKSDCPDSTLLGTFFVENHNDPR 320
DB 242 PYQNTIPGVSNPIYIYPTTRFKTTDSSSSELTQWISSVASSCDPILLTTFNFEVHNDER 301
QY 321 FASYNIDIALAKNVAFTIINDGIPITYAGQFQHYAGNDPANDPRAVATLSGYPTDSLEYK 380
DB 302 FASMTSDGLSLNAIAFAVLLGDGIPVIYIYGOGLSGKSDPNRREALLSGYNKESDYYK 361
QY 381 LIASNAIRNVAISKDTGFVTYKKNPIYKDDTTIAMRGGTGDSQIVTILLSKSGASGSYT 440

```

```

DB 362 LIAKANAARNAVYDSSVATSQLSVIRFSDHVIATKRG-----VSVFNMKSSGSS-D 416
QY 441 LSLSGAGTATAGQQLTEVIGCTTYVSGGNPVPVAGGLPVLVYPTKLASKICS 496
DB 417 VTSMTYSSGBDIVEVLTCSVSGSSD--LQVSTQSGQFQIFVPAK--VASDICS 468

RESULT 13
US-08-204-656B-6
? Sequence 6, Application US/08204656B
? Patent No. 5538882
? GENERAL INFORMATION:
? APPLICANT: Matsui, Ikuro
? APPLICANT: Ishikawa, Kazuhiko
? APPLICANT: Miyairi, Sachio
? APPLICANT: Honda, Koichi
? TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
? TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
? NUMBER OF SEQUENCES: 18
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
? STREET: 8110 Gatehouse Road, Suite 500 East
? CITY: Falls Church
? STATE: Virginia
? COUNTRY: U.S.A.
? ZIP: 22042
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/204,656B
? FILING DATE: 02-MAR-1994
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Weiner, Marc S.
? REGISTRATION NUMBER: 32,181
? REFERENCE/DOCKET NUMBER: 234-252P
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 205-8000
? TELEFAX: (703) 205-8050
? TELEX: 248345
? INFORMATION FOR SEQ. ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 468 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-204-656B-6

Query Match 51.7%; Score 1388; DB 1; Length 468;
Best Local Similarity 54.4%; Pred. No. 2.8e-115;
Matches 259; Conservative 80; Mismatches 127; Indels 10; Gaps 5;

QY 22 TPAD-WRSQSIYFLTDFRARTDGTATCMTADOKYCGGTGOGIITDKDYIQMGFTAI 80
DB 2 TNADKMRQSIYQIVTDFRARTDGTASCTEDRIKYGSGFQGIKKLDYIKMGFTAI 61
QY 81 WTPVTAQLPOTAYAGYHGYWODIYSLNENYGTADDLKALSLAHERGMYLMDVVA 140
DB 62 WISPVENIPNTAGYAGYHGMKMKNIYKINENFTADDLKSLAQLHRLMDLMDVIVT 121
QY 141 NHMGYDAGSSVDYSEVFEKPFSSQDYFHPFCFIONYEDQTOVEDCWLGDNTVSLPDLDTK 200
DB 122 NHYGSDDSGSDIDYSEVTEPFNDQKYFHNKCLISNDDQAOVQSCWEGSSVALPDLFRED 181
QY 201 DVKRKNEDYDWGSLVNSYSDGLRIDTVKHYQKDFWPGYNKAAGYVTCGEVLDGDPAYTC 260
DB 182 SDVASVFNKWDYFNKSIDGLRIDSAKHYDQGFPEFVSASGVYSGVEFGQDPAYTC 241

```


TELEX: 248345
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-467-831-8

Query Match 51.7%; Score 1388; DB 1; Length 468;

Best Local Similarity 54.4%; Pred. No. 2,8e-115;
Matches 259; Conservative 80; Mismatches 127; Indels 10; Gaps 5;

QY 22 TPAD-WRSQSIYFLITDRFARTDGTATCMTADQKYGCGTWGIIIDKLDYIQNGFTAI 80
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 2 TNADKWRQSQSIYQIVTDRFARTDGTASCTWEDRLYCGSFGGIKKLDYIKDWGFTAI 61
QY 81 WITPTAQLPQTAVYAGDHYGVWODIYSLNENYGTADDLKALSALHERGMYIMVDYA 140
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 62 WISPVENIPDNTAGYAHGLMKNTYKINENGTADDLKSLAQLDHRDMLMVDIYT 121
QY 141 NHMGYDAGSSVDYSVFEKPFSSODYFHPFCFTQNTEDQYEDCWLGDNVSLPDLDTTK 200
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 122 NHGSDSGSDSIDSEYFPFDQKTFHNYCLISNYDDQAQYQSCWEGDSSVALPDLRTED 181
QY 201 DVYKNEMYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYNKAGVYCIGEVLDDPAYTC 260
| : | : | : | | | | | | | | | | | : | : | : | : | : | : | : | : | : |
Db 182 SDVASVFNSWVKDFVGNYSIDGLRIDSARKHVDQGFPPDFVSAGVYSVEVFOGDPAYTC 241
QY 261 PYQNVMDGVLPYPIYYPLNLFKSTSGSMDLLYNNINTVKSDDCPDSTLLGTFVENHDNR 320
| | | : | | | | | | | : | : | : | : | : | : | : | : | : | | | | | |
Db 242 PYQNTIPGVSNPLPYPTTRFKTIDSSSELQMTISSVASSCSDPTLLTNFVENHDNR 301
QY 321 FASYTNDIALAKNVAFLIINDGIPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 380
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 302 FASMTSDQSLISMAIAFVLLGDGIPVIYGGQGLSGKSDPNNREALMLSGYNKESDYK 361
QY 381 LIASANAIRNVAISKDGFVYKKNPIYKDDTTIAMRKGTGDSQIVTILSNKGASGDSYT 440
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 362 LIAKANARNAAYVQDSSYATQSLSVIFSNHVIATKRGSS---VVSVENNIGSSGSS-D 416
QY 441 LLSGAGYTAGOOLLEVYIGCTTVYVSGDGNVPVWAGGLPRVLYPTEKLAGSKICS 496
: : : | | : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 417 VTISNTGISSEDELVEYLTCTVSGSSD--LQVSIQGGQPOLIVPAK--YASDICS 468

Search completed: June 13, 2002, 08:54:29
Job time: 314 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:40 ; Search time 38.1 Seconds
(without alignments)
506.098 Million cell updates/sec

Title: US-09-710-339-2

Perfect score: 2684

Sequence: 1 MVAWMSILFLYGLQVAPALA.....LPRVLYPTREKLAGSKICSSS 498

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2684	100.0	499	1	AMYA_ASPOR
2	2675	99.7	499	1	AMY_ASPSH
3	2674	99.6	499	1	AMYB_ASPAW
4	2662	99.2	498	1	AMYA_ASPAW
5	1778	66.2	484	1	AMYA_ASPNG
6	1449	54.0	512	1	AMY1_DEBAC
7	1397.5	52.1	494	1	AMY1_SACFI
8	1015.5	37.8	513	1	AMY2_SCHPO
9	899	33.5	581	1	AMY1_SCHPO
10	898	33.5	478	1	YDDL_SCHPO
11	670	25.0	1196	1	AMYB_PAPPO
12	456.5	17.0	713	1	CDGT_BACSO
13	454.5	16.9	712	1	CDGT_BACSO
14	444.5	16.6	713	1	CDGT_BACSP
15	440.5	16.4	528	1	AMY_BACCI
16	433	16.1	710	1	CDGT_THETU
17	428.5	15.9	655	1	CDGT_KLEPN
18	428	15.9	703	1	CDGT_BACSO
19	424.5	15.8	919	1	AMY_STRLI
20	418	15.6	713	1	CDGT_BACCI
21	418	15.6	714	1	CDGT_BACCI
22	414.5	15.4	713	1	CDGT_PAEWA
23	412	15.4	717	1	AMYR_BACSO
24	407	15.2	704	1	CDGT_BACOH
25	406.5	15.1	718	1	CDGT_BACCI
26	405.5	15.1	718	1	CDGT_BACSO
27	405	15.1	713	1	CDGT_BACSO
28	399.5	14.9	718	1	CDGT_PAEWA
29	385.5	14.4	711	1	CDGT_BACST
30	368	13.7	588	1	NEPU_BACST
31	337	12.6	586	1	AMY1_BACAD
32	328.5	12.2	676	1	AMY1_ECOTI
33	317.5	11.8	591	1	CDAS_BACSH

34	315	11.7	562	1	AMY2_DICTH	P14898 dictyoglomu
35	295.5	11.0	574	1	CDAS_THETU	P29964 thermococci
36	270	10.1	498	1	AMY2_DICTH	P14899 dictyoglomu
37	269	10.0	585	1	NEPU_THETU	P08751 thermococci
38	263	9.8	1481	1	APU_THETU	P38939 t_amiylopull
39	261	9.7	471	1	AMY_TENMO	P56634 tenebrio mo
40	259.5	9.7	518	1	AMY6_BACSO	P19571 bacillus sp
41	257	9.6	489	1	AMY_TRICA	P09107 tribolium c
42	255.5	9.5	561	1	TREC_BACSO	P39795 bacillus su
43	253.5	9.4	520	1	AMY_BACME	P20845 bacillus me
44	252.5	9.4	1475	1	APU_THETU	P16950 t_amiylopull
45	247	9.2	478	1	AMY_STRLX	P08486 streptomyce

ALIGNMENTS

RESULT 1
ID AMYA_ASPOR STANDARD: PRT: 499 AA.
AC P10529: P11763: Q00250:
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
DE (1,4-alpha-D-glucanohydrolase).
GN AMY1 AND AMY2 AND AMY3.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 63303;
RX MEDLINE=89237897; PubMed=2785629;
RA Wirsel S., Lachmund A., Wildhardt G., Rutkowski E.;
RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
RT intron-exon organization.";
RL Mol. Microbiol. 3:3-14(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89378767; PubMed=2789162;
RA Gesner M.J., Dove M.J., Seliy V.L.;
RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
RT containing eight introns.";
RL Gene 79:107-117(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
RT of Aspergillus oryzae.";
RL Agric. Biol. Chem. 53:593-599(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=90128276; PubMed=2612911;
RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
RA Uda K. S.;
RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
RT evidence for multiple related genes.";
RL Gene 84:319-327(1989).
RN [5]
RP SEQUENCE OF 22-499.
RA Toda H., Kondo K., Narita K.;
RT "The complete amino acid sequence of Taka-amylase A";
RT Proc. Jpn. Acad. B, Phys. Biol. Sci. 58:208-212(1982).
RN [6]
RP SEQUENCE OF 206-225.
RX MEDLINE=74001521; PubMed=4733850;
RA Isemura S., Ikenaka T.;
RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
RT A with trypsin and chymotrypsin.";
RL J. Biochem. 74:1-10(1973).
RN [7]

DE Alpha-amylase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
glucanohydrolase B).
GN AMB.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
ON NCBI_TaxID=105351;
RX MEDLINE=90254827; PubMed=2340591;
RA Kodama D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52756; CAA36967.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 499 ALPHA-AMYLASE B.
FT ACT_SITE 227 227 BY SIMILARITY.
FT ACT_SITE 231 231 BY SIMILARITY.
FT ACT_SITE 251 251 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 261 304 BY SIMILARITY.
FT DISULFID 461 496 BY SIMILARITY.
FT CARBOHYD 218 218 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;
Query Match 99.6%; Score 2674; DB 1; Length 499;
Best Local Similarity 99.6%; Pred. No. 5.2e-176;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 MWAMMSLFYGLGVAAAPALAAFPADWRSQSYFLITDFARFDTSTATCATCTADQKCGG 60
DB 2 MWAMMSLFYGLGVAAAPALAAFPADWRSQSYFLITDFARFDTSTATCATCTADQKCGG 61
OY 61 TWGIGIIDLKIDYIQGMGFPAIWTPTVTAOLPOTAYGDAHYGYWOODIYSLNENGTADDL 120
DB 62 TWGIGIIDLKIDYIQGMGFPAIWTPTVTAOLPOTAYGDAHYGYWOODIYSLNENGTADDL 121
OY 121 KALSSALHERGMVLMVDVYANMGYDAGSSVDYVFPEFSSODTFHFECFIONEDQTO 180
DB 122 KALSSALHERGMVLMVDVYANMGYDAGSSVDYVFPEFSSODTFHFECFIONEDQTO 181
OY 181 VEDCWLGNVTYSLPDLDTTKDYVKNENYDWGSLVSNYSIDGLRIDYKHYQKDFWPGYN 240
DB 182 VEDCWLGNVTYSLPDLDTTKDYVKNENYDWGSLVSNYSIDGLRIDYKHYQKDFWPGYN 241
OY 241 KAAGVYCIGEVLDDGPAYTCYQNVMDGVLPYIYPLLNAFKSTSGSMDLYNMINTVK 300

DB 242 KAAGVYCIGEVLDDGPAYTCYQNVMDGVLPYIYPLLNAFKSTSGSMDLYNMINTVK 301
OY 301 SDCPDSTLLGTVEVHNDRFASYNNDIALAKNVAFTIILNDGJPIIYAGOEQYAGSND 360
DB 302 SDCPDSTLLGTVEVHNDRFASYNNDIALAKNVAFTIILNDGJPIIYAGOEQYAGSND 361
OY 361 PANREATWLSGYPTDSELYKLJASANAIRNVAISKDTGEVYTKMNPYKDDTTIAMRGT 420
DB 362 PANREATWLSGYPTDSELYKLJASANAIRNVAISKDTGEVYTKMNPYKDDTTIAMRGT 421
OY 421 DGSQIYTIISNKGASDSTTSLSGAGCTAGQOLTEVIGCTTVYGSNDPVPWAGSLP 480
DB 422 DGSQIYTIISNKGASDSTTSLSGAGCTAGQOLTEVIGCTTVYGSNDPVPWAGSLP 481
OY 481 RVLYPTEKLKSGKICSS 498
DB 482 RVLYPTEKLKSGKICSS 499
RESULT 4
ID AMYA.ASPAW STANDARD; PRT; 498 AA.
AC 002905;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
glucanohydrolase A).
GN AMYA.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
ON NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Kodama D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52755; CAA36966.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 498 ALPHA-AMYLASE A.
FT ACT_SITE 227 227 BY SIMILARITY.
FT ACT_SITE 231 231 BY SIMILARITY.
FT ACT_SITE 251 251 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 261 304 BY SIMILARITY.
FT DISULFID 461 496 BY SIMILARITY.

FT CARBOHYD 218 218 N-LINKED (GLCMAC. . .) (POTENTIAL).
SQ SEQUENCE 498 AA; 54880 MM; 765851IBC01A8A01 CRC64;

Query Match 99.2%; Score 2662; DB 1; Length 498;
Best Local Similarity 99.6%; Pred. No. 3.4e-175;
Matches 493; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVAMWSLEFLVGLVAAAPALATPADMRSQSIYELTDREAFRTDGTATCATCNADOKYCGG 60
DB 2 MVAMWSLEFLVGLVAAAPALATPADMRSQSIYELTDREAFRTDGTATCATCNADOKYCGG 61
QY 61 TWGGIDKLDYIOGMGFTAIWTPVTAQLPOTATAGAHYGYWQODIYSLNENYGTADDL 120
DB 62 TWGGIDKLDYIOGMGFTAIWTPVTAQLPOTATAGAHYGYWQODIYSLNENYGTADDL 121
QY 121 KALSALHERGMVLMVAVNHNMGVDGAGSSVDYSVFKPPSSODYFHPFCIONYEDQTO 180
DB 122 KALSALHERGMVLMVAVNHNMGVDGAGSSVDYSVFKPPSSODYFHPFCIONYEDQTO 181
QY 181 VEDCMLGDNVSLPDLDTTKDVAKENMYDMVGLSVSNYSIDGLRIDTVKHKQDFWPGTN 240
DB 182 VEDCMLGDNVSLPDLDTTKDVAKENMYDMVGLSVSNYSIDGLRIDTVKHKQDFWPGTN 241
QY 241 KAAGYTCIGEVLDGDPATYCPQNVMDGVNLPIYPLNAFKSTSGMDLXNMTNFK 300
DB 242 KAAGYTCIGEVLDGDPATYCPQNVMDGVNLPIYPLNAFKSTSGMDLXNMTNFK 301
QY 301 SDPCDSTLLGTFEVENHNDPRFASVTNDIALAKNVAFFIILNDGPIIYAGOEQHYAGND 360
DB 302 SDPCDSTLLGTFEVENHNDPRFASVTNDIALAKNVAFFIILNDGPIIYAGOEQHYAGND 361
QY 361 PANRATATLSCYPRDSELYKLIASANAIRNVAISKOTGFYTKNMPYKXDDTTIARKEG 420
DB 362 PANRATATLSCYPRDSELYKLIASANAIRNVAISKOTGFYTKNMPYKXDDTTIARKEG 421
QY 421 DGSQIVTILSNKKGASGDSYTLSELGAGYTAGOQLTEYIGCTTVYVSGDGNVPYPMAGLP 480
DB 422 DGSQIVTILSNKKGASGDSYTLSELGAGYTAGOQLTEYIGCTTVYVSGDGNVPYPMAGLP 481
QY 481 RVLVYTERKLASGKIC 495
DB 482 RVLVYTERKLASGKIC 496

RESULT 5

AMYA_ASPNG STANDARD; PRT; 484 AA.
AC P36271.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Acid alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eutotiomyces;
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN (1)
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=91002514; PubMed=2207069;
RA Boel E., Brady L., Brzozowski A.M., Derewenda Z., Dodson G.G.,
Jensen V.J., Petersen S.B., Swift H., Thim L., Woldike H.F.,
"Calcium binding in alpha-amylases: an X-ray diffraction study at
2.1-A resolution of two enzymes from Aspergillus";
RT Biochemistry 29:6244-6249(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
linkages in oligosaccharides and polysaccharides.
CC -1- COPACOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR PDB: 2AAA; 15-JUL-93.
DR InterPro: IPR00461; Alpha_amylase.
DR Pfam: PF00128; alpha-amylase; 1.

KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KM Glycoprotein; 3D-structure.

FT ACT_SITE 206 206
FT ACT_SITE 210 210
FT ACT_SITE 230 230
FT ACT_SITE 297 297
FT DISULFID 30 38
FT DISULFID 150 164
FT DISULFID 240 283
FT DISULFID 440 475
FT CARBOHYD 24 24
FT CARBOHYD 157 157
FT CARBOHYD 197 197
SQ SEQUENCE 484 AA; 52935 MM; 04D596B34680656D CRC64;

Query Match 66.2%; Score 1778; DB 1; Length 484;
Best Local Similarity 66.9%; Pred. No. 1.2e-114;
Matches 318; Conservative 63; Mismatches 94; Indels 0; Gaps 0;

QY 24 ADMRSQSIYFLTDREAFRTDGTATCATCNADOKYCGGTTWGGIDKLDYIOGMGFTAIW 83
DB 4 ASMRQSIYFLTDREAFRTDGTATCATCNADOKYCGGTTWGGIDKLDYIOGMGFTAIW 83
QY 84 PVTAQLPOTATAGAHYGYWQODIYSLNENYGTADDLKALSALHERGMVLMVAVNHN 143
DB 64 PITDLPOTDAGAHYGYWQODIYSLNENYGTADDLKALSALHERGMVLMVAVNHN 123
QY 144 GYDGAQSSVDYSVFKPPSSODYFHPFCIONYEDQTOVEDCMLGDNVSLPDLDTTKDV 203
DB 124 GYDGAQSSVDYSVFKPPSSODYFHPFCIONYEDQTOVEDCMLGDNVSLPDLDTTKDV 183
QY 204 KNEWDMWGLSVSNYSIDGLRIDTVKHKQDFWPGTNKAAGYTCIGEVLDGDPATYCP 263
DB 184 KNEWDMWGLSVSNYSIDGLRIDTVKHKQDFWPGTNKAAGYTCIGEVLDGDPATYCP 243
QY 264 NVMDGVNLPIYPLNAFKSTSGMDLXNMTNFKSDPCDSTLLGTFEVENHNDPRFAS 323
DB 244 KVLDDVNLPIYPLNAFKSTSGMDLXNMTNFKSDPCDSTLLGTFEVENHNDPRFAS 303
QY 324 YTNDAALAKNVAFFIILNDGPIIYAGOEQHYAGNDPANRATATLSCYPRDSELYK 383
DB 304 YTSDSQAKNVAFFIILNDGPIIYAGOEQHYAGNDPANRATATLSCYPRDSELYK 363
QY 384 SANAIRNVAISKOTGFYTKNMPYKXDDTTIARKEGDSQIVTILSNKKGASGDSYTL 443
DB 364 TTNAIRKLIADASAYITITANDAFYTDSTNTIYAKGTSQIVTILSNKKGASGDSY 423
QY 444 SCAGYTAGOQLTEYIGCTTVYVSGDGNVPYPMAGLPYKXDDTTIARKEGDSY 498
DB 424 SCAGYTAGOQLTEYIGCTTVYVSGDGNVPYPMAGLPYKXDDTTIARKEGDSY 478

RESULT 6

AMY1_DEBOC STANDARD; PRT; 512 AA.
AC P19269;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase 1 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
glucanohydrolase).
GN AMY1.
OS Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=27300;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 26076;
RX MEDLINE=90032659; PubMed=2806251;
RA Strasser A.W.M., Selk R., Dohmen R.J., Niemann T., Bielefeld M.,
Seebach P., Tu G., Hollenberg C.P.;


```
RP SEQUENCE FROM N.A.  
RC SIRAIn=972;  
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;  
RL Submitted (Aug-6-1997) to the EMBL/Genbank/DBJ databases.  
CC  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
CC linkages in oligosaccharides and polysaccharides.  
CC  
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY  
CC SIMILARITY).  
CC  
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
DR EMBL; Z98762; CAB11471.1; -  
DR HSSP; P10529; 7TAHA.  
DR InterPro; IPR000461; Alpha-amylase.  
DR Pfam; PF00128; alpha-amylase; 1.  
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;  
KW Calcium; Glycoprotein; Signal.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 1 513 AA; PROBABLE ALPHA-AMYLASE C4A8.01.  
FT ACT_SITE 226 513 BY SIMILARITY.  
FT ACT_SITE 226 226 BY SIMILARITY.  
FT ACT_SITE 230 230 BY SIMILARITY.  
FT ACT_SITE 250 250 BY SIMILARITY.  
FT ACT_SITE 318 318 BY SIMILARITY.  
FT DISULFID 52 60 BY SIMILARITY.  
FT DISULFID 171 184 BY SIMILARITY.  
FT DISULFID 260 304 BY SIMILARITY.  
FT DISULFID 434 488 BY SIMILARITY.  
FT CARBOHYD 162 162 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 357 357 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 513 AA; 58715 MW; 455DD97EFA428C182 CRC64;
```

Db 417 MWMTKG-----PVITFTNYGAVDKEYLIKMGFS-----ETMIDLTCTLIEV--EGEVMR 465
QY 473 VPMAGLPRVLYP 485
Db 466 TSIRKGEKRLYP 478

RESULT 9
AMY1_SCHPO STANDARD; PRT; 581 AA.
ID AMY1_SCHPO 009840:
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE Probable alpha-amylase C23D3.14C precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
GN SPAC23D3.14C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Nidbett D., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.,
RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: 264354; CAA91249.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR00461; Alpha-amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KW Calcium; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 581 PROBABLE ALPHA-AMYLASE C23D3.14C.
FT ACT_SITE 233 233 BY SIMILARITY.
FT ACT_SITE 237 237 BY SIMILARITY.
FT ACT_SITE 325 325 BY SIMILARITY.
FT DISULFID 56 64 BY SIMILARITY.
FT DISULFID 176 191 BY SIMILARITY.
FT DISULFID 267 311 BY SIMILARITY.
FT CARBOHD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 581 AA; 67004 MW; FE9DE9D323E1890 CRC64;

Query Match 33.5%; Score 899; DB 1; Length 581;
Best Local Similarity 36.2%; Pred No. 2.4e-54;
Matches 180; Conservative 94; Mismatches 207; Indels 16; Gaps 8;

QY 2 VAWMSLEFLGYGVAPALATPA-DWRSQSIYELLDRFARIDGST-TATCNTADQKYCG 59
Db 11 IGMMLLFAF-----IPAYAGSHAEKMKRSIYGIITDRSLLEGATERPCDPVPMVCG 65
QY 60 GTWQGIIDKLDYIQMGFLTAITPTTAQLPQTAYGADYHOYWOODIYSLNENGTAD 119
Db 66 GTWNGIRNHLDYIQMGFLTAITPTTAQLPQTAYGADYHOYWOODIYSLNENGTAD 125

QY 120 LKALSALHERGYLMVDVYANMGVDGAGSSVDYSVFPPSSODYFHPFCFIONYE--D 177
Db 126 FMEILQELHKKRDWILLDAVAINSMALINGLEQMSSEKVIYFPDASFPHHCWV-DYESND 184
QY 178 QTVQEDCWLGMDNVSLPDDITTKDVKYKMEYDWGSLVSNYSIDGLRIDYKHKVOKDFWP 237
Db 185 IESVQNCWLGDEHLLADVDTENEVVLVLEKMKVQVEYDIDGIRFDAIKHAIIEFWL 244
QY 238 GYNKAAGVYCIGEVLDGDPAITCPYONV-MDGLVNYPIYPLLMFKPSSGMDLYMI 296
Db 245 RMSKADIFFTIGEFYFGSPAEADYONSGIDSEFLNPEPLTWPMFNNGLQCELAIAI 304
QY 297 NTFKSDCPSTLLGTFVEHNDNPRFASYNTDIALAKNVAFTIINDGPIIYAGOEONYA 356
Db 305 NQINEECNDINVIQGTIGNHDLPRISHNNTDQARINATTFMMMDGPIIYIGTEQNFN 364
QY 357 GGNDPANREATWLSGYPTDSELYKLIAANAIKNAISKDTGCVYTKKMPYTKDDTTAM 416
Db 365 SYHDPFNREALMWSNFMENYKKLIGILNRFKRSYOROEYVNTNSTILSVKIHHTIV 424
QY 417 RKGTDSQIYITLISNKGASGDS-YTSLSGAGYTAQQLTEVYIGCTTVYSGDNVPM 475
Db 425 QK---LNVITVLNNGYIHNEERLSIVKPLGASPKOTFEFDIINOKYVNTDGLKVI 480
QY 476 AGGLPRVLYPTEKLAGS 492
Db 481 TNGFPVLYPTSKIRTS 497

RESULT 10
YDYL_SCHPO STANDARD; PRT; 478 AA.
ID YDYL_SCHPO 010427; Q9JUN0:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 55.4 kDa protein C188.01c in chromosome III.
GN SPC1188.01C OR SPC11810.09C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Ramsperger U., Pohl T., Wood V., Rajandream M.A., Barrell B.G.,
RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AL121783; CAB57851.1; -.
DR HSSP; AL049662; CAB41221.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR00461; Alpha-amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase.
FT ACT_SITE 216 216 BY SIMILARITY.
FT ACT_SITE 220 220 BY SIMILARITY.
FT ACT_SITE 308 308 BY SIMILARITY.
FT SEQUENCE 478 AA; 55442 MW; 3C65EAD005DF73E CRC64;

Query Match 33.5%; Score 898; DB 1; Length 478;

Best Local Similarity 39.18; Pred. No. 2.2e-54;
Matches 182; Conservative 87; Mismatches 184; Indels 12; Gaps 7

QY	26	WRSQSYTEFLDREAFRTGSGTATCNCADQKCGGSGWQGIIDKLDIYGMEFTAIWTPV	85
Db	19	WRSQSYTEFLDREAFRTGSGTATCNCADQKCGGSGWQGIIDKLDIYGMEFTAIWTPV	76
QY	86	TAQLPQTAYGDAYGHYWOODIYSLENBNGYATADCLKALSALHERGMYLWVYVANHGY	145
Db	77	VKNISITDCCGAYGWAQMDOTLNMENFCEDEDLKELVNAIHERKMLCMYDVIYNHGH	136
QY	146	DGAGSSVDYSVKRPSSDDYHFPFCFIONEDPTGYEDCDMDGNVSLPDDPTTKRDYKN	205
Db	137	AGS-KRYVNLFLKRPNSCKRYHNMFPQONTDDHEITITGMDSDHNPLDIRTEKNEYRK	195
QY	206	EWDYWDGSLVSNYSTDLGRIDTYKHVQKDFWPGYNKAAGYCYCISGYLDGDPAYTCPTONY	265
Db	196	FQGNWVSDIKRKYODGFRDLRTAKHVEKSFPEFIEANAVFTTGEVFEFGDPRVVGDOXY	255
QY	266	MDVLVPIYYLYPLNAFKSTSGSMDDLYN-MINTVSDCPDSTLLGTFEVNHNDPRASY	324
Db	256	LPSTLMEFLPFLQRTFTFLDPKHSNFSFYDKAAVLDVNHFKVDYTVLCNPLENDDPRFHE	315
QY	325	TNDIALAKVNAAPFLINDGIFPIYAGQDQHYAGANDPANREXTWLSGPTDSBLKLIAS	384
Db	316	TKDIALALMALTALFPMQIDPIIYGGQDHFDDGSDPNREGIMSKSKYNTSPMERHLIS	375
QY	385	ANAIRNYATSKDTGVTGNKMPYIKDDTT-IAMRKGTDSQJVTLLISNKA-SGDSTYLS	442
Db	376	MIRTPONLETPPEFTYVLISFQLYIDDSYVFTFRPG----YIAISNGSPTSPKVEID	430
QY	443	LSAGYTAGQGLEVIIGCTTGVGSDGNVNPVMAAGLPRVYLPTE	487
Db	431	LKHAKVEVPSSFTDILTKTITPC-KDKHLIKKSGSLPKILISSD	474
RESULT	11		
AMYB_PAPEO			
ID	AMYB_PAPEO	STANDARD:	PRT; 1196 AA.
AC	P21543:		
DT	01-MAY-1991 (rel. 18, Created)		
DT	01-MAY-1991 (rel. 18, Last sequence update)		
DT	01-MAR-2002 (rel. 41, Last annotation update)		
DE	Beta/alpha-amyase precursor [includes: Beta-amyase (EC 3.2.1.2);		
DE	Alpha-amyase (EC 3.2.1.1)].		
OS	Paenibacillus polymyxa (Bacillus polymyxa).		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
OC	Bacillus/Staphylococcus group; Paenibacillus.		
OX	NCBI_Taxid=1406;		
RN	[1]		
RN	[2]		
RP	SEQUENCE OF 689-1196 FROM N.A., AND PARTIAL SEQUENCE.		
RC	SPRAIN-72;		
KX	MEDLINE=89123046; PubMed=2464578;		
RA	Uozumi N., Sakurai K., Sasaki T., Takekawa S., Yamagata H.,		
RA	Tsukagoshi N., Uda K. S.;		
RT	"A single gene directs synthesis of a precursor protein with beta-		
RT	and alpha-amyase activities in Bacillus polymyxa."		
RL	J. Bacteriol. 171:375-382(1989).		
RN	[3]		
RP	SEQUENCE OF 1-776 FROM N.A.		
RC	SPRAIN-ATCC 8523;		
KX	MEDLINE=87231094; PubMed=2438660;		
RA	Rhodes C., Strasser J., Friedberg F.;		
RT	"Sequence of an active fragment of B. polymyxa beta amyase."		

FT CONFLICT 736 736 N -> K (IN REF. 3).
 FT CONFLICT 741 741 N -> S (IN REF. 3).
 FT CONFLICT 758 758 S -> N (IN REF. 3).
 SQ SEQUENCE 1196 AA; 130893 MW; A41EA6B70FE257064 CRC64;

Query Match 25.0%; Score 670; DB 1; Length 1196;
 Best Local Similarity 34.2%; Pred. No. 3e-38;
 Matches 172; Conservative 82; Mismatches 177; Indels 72; Gaps 17;

QY 13 QVAPALAAP-----ADMRSGSIYFLIDRFARNDGSGT-----ATCMTADO 55
 DB 718 QTEPSPGAFPSMDGLTSLNPTSMTDWSKOSITIFIMDRSNGDPSNDYGGNSNSDQ 777
 QY 56 -KYCGWQODIYSLNENYGTADIKALSSALHERGMYLVANHMGYDAGSSVDYSVKRPSSODYFHFECI-- 114
 DB 778 RKMHGDPGQGIINKLDYIKMNGFAIMITPVTKMSEY-----AYHGHTYDEYAVDGLH 832
 QY 115 GTADDLKALSSALHERGMYLVANHMGYDAGSSVDYSVKRPSSODYFHFECI-- 172
 DB 833 GTMDKLELYRKADKNIAYVAVVNVHNTGDFOPNGFARA--PFDKADWYHHNGDITD 889
 QY 173 --QNVETQVEDCWLGNFVSLPDLDTKDYVKNEMYDWGSLVSNYSIDGLRIDTVKH 230
 DB 890 GDYSSNMWKTEN---GD-VAGLDDLDHNEPATANELKNIMKWLNETGIDGLRIDTVKH 945
 QY 221 VQKDFWPGYKNAGVYICGEVLDDPAYTCYQNVMDGVLYNPIYPLNANF-----K 283
 DB 946 VPKGFLKFDQAMTFMGELFHFHGDPAVYGDYTRFLDAALDFPMXYTIKDFVGHDSMRK 1005
 QY 284 STSGMDLYMNTVKSDCDSTLLGTEVENHNHNPFR-----ASYTNDIALAKNV 334
 DB 1006 IKDRYSDDRYR-----DAQTNGVFDNHVKRFRLANDASKPGANNDKMQOL-KAA 1055
 QY 335 AAFILNDGIPILIYAGOEYHAGNDPANREATWLSGYPDSELYKLASANAIRNVAIS 394
 DB 1056 LGFTLTSRGIFILTYOGTEGYSGGDDPANREN---MNFNANHDLXYQYIAKINYRNNHRA 1112
 QY 395 KDTGVTYKKNPIYKDDTTIMARKGTGSOQVITLISNKGAGSDSYTSLSG-AGYTAGOO 453
 DB 1113 LQNGSQREK---WVDSFYSFQRSKNGDEAIVFLNN---SWNSQTRITGNFNLMSGTR 1165
 QY 454 LFEVIGCTTVYVGSNGNVPVPA 476
 DB 1166 LFNQLSNDSVOI-NGSITVILA 1187

RESULT 12
 CDGT_BACSO STANDARD: PRT: 713 AA.
 ID CDGT_BACSO STANDARD: PRT: 713 AA.
 AC P05618:
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Cyclomaltodextrin glucanotransferase precursor (EC 2.4.1.19)
 DE (Cyclodextrin-glycosyltransferase) (CGTase).
 GN CGT.
 OS Bacillus sp. (strain 1011).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87308036; PubMed=2957361;
 RA Kimura K., Kataoka S., Ishii Y., Takano T., Yamane K.;
 RT "Nucleotide sequence of the beta-cyclodextrin glucanotransferase gene
 of alkaliophilic Bacillus sp. strain 1011 and similarity of its amino
 acid sequence to those of alpha-amylases.";
 RL J. Bacteriol. 169:4399-4402(1987).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RA Harata K., Haga K., Nakamura A., Aoyagi M., Yamane K.;
 RT "X-ray structure of cyclodextrin glucanotransferase from alkaliophilic

RT Bacillus Sp. 1011. Comparison of two independent molecules at 1.8-A
 resolution.";
 RL Acta Crystallogr. D 52:1136-1145(1996).
 CC -1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation
 of a 1,4-alpha-D-glucosidic bond.
 CC -1- COFACTOR: BINDS TWO CALCIUM IONS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE
 IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND
 IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER
 ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN
 ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE
 MALTOLOGOSACCHARIDE PRODUCED.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M17366; AAA22308.1; .
 DR PIR: A26678; ALBSG1.
 DR PDB: 1PAM; 1JAN-97.
 DR InterPro: IPR000461; Alpha_amylase.
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR002909; IPT_TIG.
 DR pfam: PF001128; alpha-amylase; 1.
 DR pfam: PF02806; alpha-amylase_C; 1.
 DR pfam: PF00686; CBD_4; 1.
 DR pfam: PF01833; TIG; 1.
 DR ProDom: PD001568; CBD_4; 1.
 DR Transfaserase; Glycosyltransferase; Calcium; Signal; 3D-structure.
 KW TRANSFASERASE; 1
 FT SIGNAL 27
 FT CHAIN 28 713
 FT DOMAIN 28 165
 FT DOMAIN 166 229
 FT DOMAIN 230 433
 FT DOMAIN 434 522
 FT DOMAIN 523 609
 FT DOMAIN 610 713
 FT DISULFID 70 77
 FT ACT_SITE 256 256
 FT ACT_SITE 284 284
 FT ACT_SITE 355 355
 FT SEQUENCE 713 AA; 78340 MW; 524B259526B56C52 CRC64;

Query Match 17.0%; Score 456.5; DB 1; Length 713;
 Best Local Similarity 28.3%; Pred. No. 6.5e-24;
 Matches 149; Conservative 73; Mismatches 216; Indels 89; Gaps 21;

QY 1 MVAWSLFL-YGLQVAPALAA-----TPADMRSGSIYFLIDRF-----ART 42
 DB 7 LTAWTLMLSLTGLSPVHAAPDTSVSKNKFSTVYITQFTDRSDGNPANNPAGALF 66
 QY 43 DGSFTATCNTADQKYGCGTQGIIDKLD--YIOGSGFTAITWTPYTAOLPQTTAY---G 96
 DB 67 DGSCTNL-----RLYCGGDMOGIINKINDGYLNGMTAITWISQPEVNISVYNSGVNN 121
 QY 97 DAYHGWQODIYSLNENYGTADIKALSSALHERGMYLVANHMGYDAGSSVDYSVKRPSSODYFHFECI-- 144
 DB 122 TAYHGYWARDFEKKNPAYGTMODFKNLIDTAHMHNIKVIIDFAPNHTSPASSDDPSFAEN 181
 QY 145 ---YDAGSSVDYSVKRPSSODYFHF-----PCEFTON--YEDOTQVEDCWLGDNTVSLP 194
 DB 182 GRLYDNGNLGGYT---NDTONLFHHYGGTDSITENGICYKNIYDLADLNHNKSSVDY- 236
 QY 195 DLDFTKDYVKNEMYDWGSLVSNYSIDGLRIDTVKHV---QKDFWPGYKNAGVYICGE 250

FT ACT_SITE 357 357 BY SIMILARITY.
SQ SEQUENCE 528 AA; 57940 MW; 455740202FA29753 CRC64;

Query Match 16.4%; Score 440.5; DB 1; Length 528;
Best Local Similarity 27.1%; Pred. No. 5.5e-23;
Matches 149; Conservative 78; Mismatches 206; Indels 117; Gaps 25;

```
QY 1 MYAMMSLFYGLQVAPALATPA---DWRSQSIYFLIDRFAPRDSST-----47
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 IIALIAIAFGSVAPEAAAPATSVSKONFSTDVYIQTDRF--VDGNTANNPAGSAYD 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 48 ATCNTADCKYCGGTMOGIDKLD--YQGMGFATITVTYATLQPTAY----GDAYHG 101
   ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : :
Db 69 ATCSTNLKLYCGDMQGMINKINDGYFTGMTALMISOPVENTYSVINSGVNNFTAYHG 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 102 YWODIYSLNENYGTADDLKALSSALHERGMVLMVDVANHM-----143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 YWADFCKTNPAFGSMTDFANLISAHSRNKIVVIDFAPRHTSPAMETNASFGENGKLYD 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 144 -----GYDGAASSVDYSVEKPFSSQDYFH-----PFCFIQNTEDQTYVEDCWLGDNTYS 192
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 NGTLIGSTG-----DFNGYFHNGSGTDFSTLKN-----GIYKNLYD 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 193 LPDLDTTRDYY---KNEMYDWGSLVSNYSIDGLRIDTVKHY---QKDFWPGYNKAAG 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 LADLNHNKSTIDTFEKNKAIKRLML---DMGIDGIRVDVAKHMFPGMKNMMSIYSYKP 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 245 VYCIQE-----VLDGPAITCPYQNMQGVLPYIYPLNFKSTSGSMDDLYNMINT 298
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 281 VFTFGEWFLGTNETDANNITYFANESGM--SLDREFRSQAKROYFRDGSDDTMYGLDSMLS 338
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 299 VSKDCPDSTLLGTFEVHNHNPRF-ASYTNDIALAKNYAFIILNDGIPITIVAGOEHOYAG 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 339 TADVIYSVNDQVITLDNNDMDKRFVYSGANGKRL-EQALALITLSRGVPAIYGTEDYMTG 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 358 GNDPANREATWLSGYPIDSELYKLILASAMIR--NTAISKDGEFTYKKNMPLYKDDTTIA 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 NQDPNNRAK--MSSFSTSTAYVNIKLAFLRKSNPALAYGT--TQQRW--INNDVYIY 450
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 416 MRKGTGSGQIYVILSNKSGAGSDSYTL-----SLSGAGYFAGQQLTEVIGCTTIVTGSDDG- 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 451 ERK--FGNNVAVVAINKNLT--SSYSIAGLNTSLPSTGYT--DVLANSLSGNSITVGSSEA 505
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 470 -NVPIVMAAG 478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 506 VNTFTLQAGG 515
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: June 13, 2002, 09:11:48
Job time: 1028 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 13, 2002, 08:53:25 ; Search time 158.68 Seconds
(without alignments)
7.700 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160
Perfect score: 57
Sequence: 1 SSVSDYSVFKEPF 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	478	15 AAR6065	Mutant alpha-amy
2	57	100.0	478	16 AAR79025	Mature taka-amy
3	57	100.0	478	16 AAR72450	Aspergillus oryzae
4	57	100.0	478	16 AAR78270	Aspergillus oryzae
5	57	100.0	478	17 AAR14500	Aspergillus oryzae
6	57	100.0	498	22 AAB84206	Amino acid sequenc
7	45	78.9	511	9 AAP81161	Recombinant alpha-
8	45	78.9	511	9 AAP81180	Sequence of alpha-
9	45	78.9	512	11 AAR07574	Alpha-amyase enco
10	41	71.9	493	17 AAR88212	Alpha-amyase. Th
11	39	68.4	5250	21 AAY53678	Sequence gi/442661

12	39	68.4	5322	22 ABB69422	Drosophila melanog
13	38	66.7	55	20 AAY41363	Human secreted pro
14	38	66.7	413	21 AAB10970	H. tuberculata hem
15	38	66.7	413	22 AAB71082	H. tuberculata Hth
16	38	66.7	413	22 AAB71137	H. tuberculata Hth
17	38	66.7	414	21 AAB10995	H. tuberculata hem
18	38	66.7	414	22 AAB71107	H. tuberculata Hth
19	38	66.7	415	21 AAB10977	H. tuberculata hem
20	38	66.7	415	22 AAB71089	H. tuberculata Hth
21	38	66.7	415	22 AAB71145	H. tuberculata Hth
22	37	64.9	452	15 AAR63186	Variant alpha amy
23	37	64.9	468	13 AAR24136	Alpha-amyase vari
24	37	64.9	468	15 AAR63187	Variant alpha amy
25	37	64.9	468	15 AAR63184	Variant alpha amy
26	37	64.9	468	15 AAR63185	Variant alpha amy
27	37	64.9	494	8 AAP70571	Alpha-amyase gene
28	37	64.9	4820	22 ABB58592	Drosophila melanog
29	36	63.2	42	22 AAB68777	Human Immune/haema
30	36	63.2	165	21 AAG33468	Arabidopsis thalia
31	36	63.2	198	21 AAG33467	Arabidopsis thalia
32	36	63.2	266	21 AAG33466	Arabidopsis thalia
33	36	63.2	514	21 AAB42398	Human ORFX ORF2162
34	36	63.2	717	22 AAT19456	Human diagnostic a
35	35	61.4	84	22 AAB03143	Human gene 21 enco
36	35	61.4	121	22 AAB052815	Protonibacterium
37	35	61.4	162	22 AAB03117	Human gene 21 enco
38	35	61.4	255	20 AAG31833	Human foetal kidn
39	35	61.4	255	22 ABB02357	Novel human diapo
40	35	61.4	255	22 AAT02922	Human PRC polypept
41	35	61.4	255	22 AAB73684	Human oxidoreducta
42	35	61.4	349	20 AAT09298	Rice beta-glucanase
43	35	61.4	377	20 AAY09306	Rice Gns9 CDS prot
44	35	61.4	390	22 AAG67096	Transcription fact
45	35	61.4	415	21 AAG21556	Arabidopsis thalia

ALIGNMENTS

RESULT 1
ID AAR6065 standard; protein; 478 AA.
AC AAR6065;
DT 18-JUL-1994 (first entry)
XX
DE Mutant alpha-amyase.
DE
KW Methionine substitution; stability; activity; detergent;
KW dishwashing agents; liquidation agents.
XX
OS Aspergillus oryzae.
XX
PN WO9402597-A.
XX
PD 03-FEB-1994.
XX
PF 06-JUL-1993; 93WO-DK00230.
XX
PR 23-JUL-1992; 92DK-0000946.
PR 16-DEC-1992; 92DK-0001503.
PR 15-MAR-1993; 93DK-0000292.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
BI Bisgard-frantzen H, Svendsen A;
DR WPI. 1994-048855/06.
XX
PT Mutant alpha-amyase from Bacillus species comprising a
PT methionine substitution - with improved stability and activity at
PT low pH, for use in detergents, dishwashing agents and

CC pululanase. The method comprises replacement of a group in the

Db 130 ssdysvfkpf 140

RESULT 4
AAW8270
ID AAW8270 standard; Protein; 478 AA.
XX
XX AAW8270;
AC
XX
XX 17-JAN-1996 (first entry)
DE Aspergillus oryzae alpha amylase (mature protein).
XX
XX Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
KM starch; thermostable; methionine; Bacillus licheniformis;
KM Bacillus thermoliquefaciens; Bacillus steaerothermophilus;
XX Aspergillus oryzae.
XX
XX Aspergillus oryzae.
XX
XX W09521247-A1.
XX
XX 10-AUG-1995.
XX
XX 05-OCT-1994; 94WO-DK00371.
XX
XX 02-FEB-1994; 94DK-0000141.
XX
XX (NOVO) NOVO-NORDISK AS.
XX
XX Marcher D, Nilsson TE, Pedersen HH, Toft AH;
XX
XX WPI, 1995-283767/37.
XX
XX Use of an oxidation stable alpha-amylase - for simultaneous desizing
XX and bleaching or scouring of fabrics contg. starch or starch derivs.
XX
XX Disclosure; Page 25-26; 37pp; English.
XX
XX
XX Oxidation stable alpha amylases can be used for the simultaneous
CC desizing and bleaching or scouring of a fabric comprising starch or
CC starch derivatives. They exhibit a better heat stability, and
CC especially in the presence of oxidizing agents. They are obtained
CC from a parent alpha amylase by replacing one or more methionine
CC residues with any amino acid different from Cys or Met, preferably
CC leu, thr, ala, gly, ser, ile or asp. The parent alpha amylase is
CC pref. derived from a Bacillus species, although alpha amylases of
CC fungal origin can also be used. This sequence is the wild type
CC (unmodified) alpha amylase of Aspergillus oryzae.
XX
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 57; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSNDYSVFKPF 11
IIIIIIIIII
Db 130 ssdysvfkpf 140

RESULT 5
AAW14500
ID AAW14500 standard; protein; 478 AA.
XX
XX AAW14500;
AC
XX
XX 04-JUN-1997 (first entry)
DE Aspergillus oryzae alpha-amylase (mature protein).
XX
XX
XX alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;

KM Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;
KM calcium dependency; substrate binding; stability; pH optimum;
KM thermostability; cleavage; oligosaccharide substrate; dishwashing;
KM washing; detergent additive; fabric desizing; starch liquefaction;
KM sweetener; ethanol production; variant.
XX
XX Aspergillus oryzae.
OS
FH Location/Qualifiers
FH
FH Key
FH
FH Misc-difference 13..45
FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 7-23 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 33"
FT
FT
FT Misc-difference 14..40
FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 8-18
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 35"
FT
FT Misc-difference 28..42
FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 12-19 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 30"
FT
FT
FT Misc-difference 32..38
FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 14-15
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 32"
FT
FT Misc-difference 66..84
FT /label= loop 2 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 44-57 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 18"
FT
FT
FT Misc-difference 70..78
FT /label= loop 2 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 48-51
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 20"
FT
FT Misc-difference 98..210
FT /label= loop 3 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 117-185 of AAW14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 24"
FT
FT
FT Misc-difference 102..206
FT /label= loop 3 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 121-181
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 26"
FT
FT Misc-difference 121..181
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to this fragment is deleted or
FT replaced with a fragment corresponding to
FT 102-206 of AAW14499; claim 41"
FT
FT
FT Misc-difference 121..174
FT /note= "preferred region where at least one amino acid


```

RESULT 7
AAP81161
ID AAP81161 standard; protein; 511 AA.
XX
XX AAP81161;
AC
XX 12-OCT-1990 (first entry)
DT
XX Recombinant alpha-amylase.
DE
XX Alpha amylase; yeast; brewing; baking.
KW
XX Schwanniomycetes castellii ATCC 26076.
OS
XX EP257115-A.
PN
XX 02-MAR-1988.
PF
XX 21-AUG-1986; 86EP-0111586.
XX
XX 21-AUG-1986; 86EP-0111586.
PR
XX (HETB ) HEINEKEN TECHNISC BEHEER BV.
PA
XX Strasser A, Martens FB, Dohmen J, Hollenberg CP;
PI
XX WPI: 1988-057532/09.
DR
XX N-PSDB; AAN81477.
DR
XX Prodn. of amylolytic enzymes by transformed microorganisms -
PT comprises use of recombinant DNA technology to give transformant
PT having DNA from donor yeast.
PS
XX Disclosure: ; P: English.
XX
XX The sequence encoding the protein can be used to construct an
CC expression vector for the prodn. of alpha amylase in yeast cells.
CC The transformants and recombinant enzyme produced can be used in
CC fermentation processes e.g. in baking and brewing. The transformed
CC yeasts are esp. suitable for prodn. of low carbohydrate beers.
XX
XX Sequence 511 AA;
SQ

```

Query Match 78.9%; Score 45; DB 9; Length 511;
 Best Local Similarity 81.8%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 SSVDSVFKPF 11
    |||||
DB 165 ssvdysftpf 175

```

```

RESULT 8
AAP81180
ID AAP81180 standard; protein; 511 AA.
XX
XX AAP81180;
AC
XX 12-DEC-1990 (first entry)
DT
XX Sequence of alpha-amylase encoded by genomic DNA of Schwanniomycetes
DE castellii.
DE
XX Brewing; beer; breadmaking; biomass.
KW
XX Schwanniomycetes castellii.
OS
XX EP260404-A.
PN
XX 23-MAR-1988.
PD
XX
XX

```

```

PF 17-JUL-1987; 87EP-0110370.
XX
XX 17-JUL-1987; 87EP-0110370.
PR
XX
XX (HETB ) HEINEKEN TEC BEHEER NV.
PA
XX Strasser A, Martens FB, Dohmen J, Hollenberg CP;
PI
XX WPI: 1988-078794/12.
DR
XX N-PSDB; AAN81525.
DR
XX Polypeptide prodn. from cells transformed with yeast DNA -
PT esp. coding alpha amylase or glucosylase, able to convert
PT starch to ethanol
PS
XX Claim 18; Fig 2A-2C; 74pp; English.
XX
XX A cosmid library is constructed from genomic DNA of ATCC26076 and this
CC screened for inserts contg. the AA gene by ability to transform
CC S.cerevisiae GRF 18 to histidine prototrophy. One positive cosmid
CC (32kb; pyc1) was digested with EcoRI, religated and used to transform
CC E.coli JA221. Positive transformants contain pyc1-alpha, which includes
CC a 5kb EcoRI fragment contg. the gene in AAN81525.
CC
XX Sequence 511 AA;
SQ

```

Query Match 78.9%; Score 45; DB 9; Length 511;
 Best Local Similarity 81.8%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 SSVDSVFKPF 11
    |||||
DB 165 ssvdysftpf 175

```

```

RESULT 9
AAR07574
ID AAR07574 standard; protein; 512 AA.
XX
XX AAR07574;
AC
XX 02-FEB-1991 (first entry)
DT
XX Alpha-amylase encoded by the AMY1 gene.
DE
XX Schwanniomycetes yeast cells; expression cassette; alpha-amylase;
KW marker gene; regulon; signal peptide; terminator;
KW autonomously replicating sequence.
XX
XX Schwanniomycetes occidentalis.
OS
XX
XX Key Location/Qualifiers
FH Disulfide-bond 66..74
FT Disulfide-bond /note="1"
FT Disulfide-bond 186..200
FT Disulfide-bond /note="2"
FT Disulfide-bond 276..319
FT Disulfide-bond /note="3"
FT Disulfide-bond 475..510
FT Disulfide-bond /note="4"
XX
XX EP394538-A.
PN
XX 31-OCT-1990.
PD
XX 28-APR-1989; 89EP-0107780.
PF
XX 28-APR-1989; 89EP-0107780.
PR
XX (RHEI ) RHEIN BIOTECH GES.
PA
XX Hollenberg C, Strasser A;
PI

```

```

XX WPI; 1990-328670/44.
DR N-PSDB; AAQ06388.
XX
PT Transformed Schwaniumyces yeast cells - contg. an expression
PT cassette contg. regulon, DNA coding for foreign protein and
XX terminator
XX
PS Disclosure: Fig 1B(1-6); 59pp; English.
XX
CC The expression cassette may contain a regulon (R), a signal peptide
CC sequence (S), a foreign protein sequence (P) and a terminator (T)
CC of which R and/or S and/or T are derived from the alpha-amylase gene
CC of S. occidentalis.
CC R is pref. a 1.8 kb BglII-XhoI fragment (bases -1 to -540 of this
CC sequence).
CC S contains all or part of one of the sequences represented
CC in AAQ06383-87.
CC F is e.g. a cellulase, interleukin, insulin-like-growth factor,
CC interferon etc.
CC T is pref. all or part of the terminator of this sequence
CC (bp 1537-1740).
CC AMY1 acts as selective marker for the vector carrying the expression
CC cassette.
CC See also AAQ06389.
XX
SQ Sequence 512 AA:

Query Match 78.9%; Score 45; DB 11; Length 512;
Best Local Similarity 81.8%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDYSVFKPF 11
Db 166 ssvdysstfcpf 176

RESULT 10
AAR88212
ID AAR88212 standard; Protein; 493 AA.
XX
AC AAR88212;
XX
DT 03-APR-1996 (first entry)
XX
DE Alpha-amylase.
XX
KW Alpha-amylase; thermostable enzyme; baking; Aspergillus niger.
XX
Thermomyces lanuginosus CBS 224.63.
XX
FT Key Location/Qualifiers
FT Peptide 1..18
FT /label= Sig_peptide

W09601323-A1.
XX
PN 18-JAN-1996.
XX
PD 03-JUL-1995; 95WO-EP02607.
XX
PF 04-JUL-1994; 94GB-0013419.
XX
PR (DANI-) DANISCO AS.
XX
PA Michelsen B, Rasmussen P;
XX
PI WPI; 1996-087673/09.
XX
DR N-PSDB; AAT10562.
XX
PT Thermophilic alpha-amylase with activity range of 60-80 degrees C -
PT derived from Thermomyces lanuginosus, useful in the prepn. of

```

```

PT foodstuffs and bakery prods. esp. bread
XX
PS Claim 3; Page 36-38; 94pp; English.
XX
CC A thermostable alpha-amylase (AAR88212) of Thermomyces lanuginosus
CC CBS 224.63 is expressed in Aspergillus niger hosts using a gene
CC (AAT10562) isolated from a T. lanuginosus gene library. The
CC recombinant enzyme (54-60 kDa) shows optimal activity at
CC 60-70 deg and pH 5.8-6, has a pI of 3.7 and is active at 60-80
CC deg.
XX
SQ Sequence 493 AA:

Query Match 71.9%; Score 41; DB 17; Length 493;
Best Local Similarity 70.0%; Pred. No. 9.4;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDYSVFKPF 10
Db 150 ssvdysstfcp 159

RESULT 11
AAV53678
ID AAV53678 standard; Protein; 5250 AA.
XX
AC AAV53678;
XX
DT 22-FEB-2000 (first entry)
XX
DE Sequence g1/4426611/gb/AAD204501 from an alignment with protein 274.
XX
KW Mechanical stress; gene therapy; protein 274; osteoporosis; bone density;
KW bone development; g1/3328186.
XX
OS unidentified.
XX
PN W09960164-A1.
XX
PD 25-NOV-1999.
XX
PF 14-MAY-1999; 99WO-US11066.
XX
PR 15-MAY-1998; 98US-0085673.
XX
PA (QUAR-) QUARK BIOTECH INC.
XX
PI Einat P, Mor O, Skalter R, Feinstein E, Faerman A;
XX
DR WPI; 2000-053304/04.
XX
PT Identification of stress induced genes for determining risk and
PT preventing, treating or controlling osteoporosis
XX
PS Claim 32; Fig 14A-T; 308pp; English.
XX
CC The present sequence is obtained from a clustal X alignment with
CC protein 274. Protein 274 was identified using the method of the invention
CC after subjecting rat osteoblasts to mechanical stress. Expression of the
CC 608 gene was found to be upregulated by about 3-fold in cells subjected
CC to mechanical strain. The specification describes a method for the
CC identification of genes responsive to a specific mechanical stress. The
CC method comprises applying the mechanical stress to an organism (tissue
CC or cells comprising bone cells), isolating the specific cellular
CC fractions and extracting mRNA from them, and differentially analyzing the
CC mRNA in comparison with control samples. The method is used to identify
CC genes whose expression is responsive to a specific stress. The identified
CC genes are employed in determining risk associated with a physiological or
CC disease state. The risk determination methods are used for testing a
CC medicament for gene therapy. These medicaments, or genes identified by
CC the method of the invention, are used for treating, preventing or
CC controlling a physiological or disease state (especially osteoporosis or

```

CC bone density or other factors causing or contributing to osteoporosis or
CC its symptoms or other conditions involved in mechanical stress or its
CC lack. The methods can also be used for advancing research or studies in
CC bone development.

XX Sequence 5250 AA;

Query Match

Best Local Similarity 68.4%; Score 39; DB 21; Length 5250;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DYSVEKPF 11

Db 5147 dytykpf 5154

WU 12

AB69422 standard; Protein; 5322 AA.

AB69422;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 35058.

Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US09231.

23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EW;

WPI; 2001-656860/75.

N-PSDB; ABL13525.

New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

Disclosure; SEQ ID NO 35058; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins
(ABBS7737-ABBS72072).

The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

Sequence 5322 AA;

Query Match

Best Local Similarity 68.4%; Score 39; DB 22; Length 5322;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DYSVEKPF 11

Db 5219 dytykpf 5226

RESULT 13

AAV41363 standard; Protein; 55 AA.

AAV41363;

02-DEC-1999 (first entry)

Human secreted protein encoded by gene 56 clone HNGFE55.

Human; secreted protein; fusion protein; gene therapy; protein therapy;
diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
developmental abnormality; foetal deficiency; blood; allergy; renal;
immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

Homo sapiens.

WO9947540-A1.

23-SEP-1999.

18-MAR-1999; 99WO-US05804.

19-MAR-1998; 98US-0078563.

19-MAR-1998; 98US-0078566.

19-MAR-1998; 98US-0078573.

19-MAR-1998; 98US-0078574.

19-MAR-1998; 98US-0078576.

19-MAR-1998; 98US-0078577.

19-MAR-1998; 98US-0078578.

19-MAR-1998; 98US-0078579.

19-MAR-1998; 98US-0078581.

01-APR-1998; 98US-0080312.

01-APR-1998; 98US-0080313.

01-APR-1998; 98US-0080314.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;

Wel Y, Endress GA, Duan RD, Kyaw H, Ebone R, Lafleur DW;

Olsen HS, Shi Y, Moore PA;

WPI; 1999-562050/47.

N-PSDB; AA224866.

New isolated human genes, useful for diagnosis and treatment of e.g.
cancers, neurological disorders, immune diseases, inflammation or blood
disorders -

Claim 11; Page 391; 484pp; English.

This sequence represents a secreted human protein encoded by the gene
clone detailed in the descriptor line. The gene can be used to generate
fusion proteins by linking to the gene to the human immunoglobulin Fc
portion (e.g. AA224802) for increasing the stability of the fused
protein as compared to the human protein only.

The invention relates to 95 novel genes and their fragments (nucleic
acid sequences: AA224811-224907; amino acid sequences AAV41308-Y41404)
which are useful for preventing, treating or ameliorating medical
conditions e.g. by protein or gene therapy. Also, pathological
conditions can be diagnosed by determining the amount of the new
polypeptides in a sample or by determining the presence of mutations in
the new polynucleotides. Specific uses are described for each of the 95
polynucleotides, based on which tissues they are most highly expressed in

CC (see AAZ24811 for described uses).

XX Sequence 55 AA;

SO

Query Match

Best Local Similarity 66.7%; Score 38; DB 20; Length 55;
Pred. No. 2.9;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDSVFKP 10

Db 21 ssvfsvfkp 30

RESULT 14

AAB10970 AAB10970 standard; Protein: 413 AA.

AC AAB10970;

DT 16-FEB-2001 (first entry)

XX H. tuberculata hemocyanin HtH1 domain d.

KW Hemocyanin; cytosolic; virucide; antibacterial; antiparasitic;

KW immunomodulatory; antihypertensive; gene therapy; tumor; treatment;

KW infection; schistosomiasis; carcinoma; bladder; epithelium; ovary;

KW breast; bronchi; colon-rectum; hypertension; vaccine; cocaine misuse;

KW pharmaceutical carrier.

OS Halictis tuberculata.

PN WO2005192-A2.

PD 21-SEP-2000.

XX 17-MAR-2000; 2000WO-EP02410.

PR 17-MAR-1999; 99DE-1011971.

XX 20-AUG-1999; 99DE-1039578.

PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

PI Markl J, Altenheim B, Lieb B, Stiefel T;

XX WPI: 2000-587517/55.

PT New nucleic acid encoding hemocyanin, useful for gene therapy of tumors

XX and for recombinant production of fusion proteins for vaccination -

XX Claim 21; Page 91-93; 163pp; German.

CC This invention describes a novel nucleic acid (I) containing a sequence

CC that encodes hemocyanin (II), a domain of (I) or its fragment with the

CC immunological properties of at least one domain of (II). The products of

CC the invention have cytosolic, virucide, antibacterial, antiparasitic,

CC immunomodulatory and antihypertensive activity. (I), and constructs

CC additionally containing antigen-encoding sequences, are useful in gene

CC therapy of tumors. Polypeptides encoded by (I) are useful for treating

CC parasitic or viral infections and tumors, particularly schistosomiasis

CC and carcinoma (of bladder, epithelium, ovary, breast, bronchi or

CC colon-rectum), also hypertension, as vaccines, for treating cocaine

CC misuse and very generally as carriers for pharmaceuticals, e.g.

CC cytosolics. They may also be used to generate antibodies (Ab). Probes

CC based on (I) and Ab are useful for detecting tumor-specific DNA in a cell

CC (by detecting specific binding to cellular DNA or proteins), particularly

CC where associated with the types of carcinoma listed above. Hemocyanins

CC can be produced recombinantly, relatively inexpensively and in adequate

CC amounts, eliminating the need to culture gastropods. When used as a

CC carrier, (II) significantly increases the half-life of the attached

CC pharmaceutical, by inhibiting ultrafiltration in the kidneys.

XX Sequence 413 AA;

Query Match 66.7%; Score 38; DB 21; Length 413;

Best Local Similarity 70.0%; Pred. No. 30;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSVDSVFKP 10

Db 200 sldysatdp 209

RESULT 15

AAB71082 AAB71082 standard; Protein: 413 AA.

AC AAB71082;

DT 24-AUG-2001 (first entry)

XX H. tuberculata HtH1 domain d.

KW HtH1; KtH1; hemocyanin; cytosolic; vasotropic; vaccine; gene therapy;

KW tumor; parasite infection; viral infection; antitumor agent;

KW abnormal blood pressure; bladder carcinoma; epithelial carcinoma;

KW ovarian carcinoma; mammary carcinoma; bronchial carcinoma;

OS Halictis tuberculata.

PN WO200114536-A2.

PD 01-MAR-2001.

XX 21-AUG-2000; 2000WO-EP08129.

PR 20-AUG-1999; 99DE-1039578.

PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

PI Markl J, Altenheim B, Lieb B, Stiefel T;

XX WPI: 2001-191646/19.

PT New nucleic acid encoding hemocyanin, useful for gene therapy of tumors

XX and for recombinant production of fusion proteins for vaccination -

XX Claim 25; Page 108-110; 180pp; German.

CC This invention describes a novel nucleic acid molecule (N1) encoding

CC hemocyanin, a hemocyanin domain, or a functional fragment of it with

CC the immunological properties of at least one domain of hemocyanin, and

CC which comprises at least one intron is new. The products of the invention

CC have cytostatic and vasotropic activity and can be used in vaccines or

CC for gene therapy. Compositions comprising nucleic acids of the invention

CC are useful in treating tumors. Compositions comprising hemocyanin

CC polypeptide are useful in treating parasite and viral infections and as

CC an antitumor agent. It is also useful in treating abnormal blood

CC pressure, bladder carcinomas, epithelial carcinomas, ovarian carcinomas,

CC mammary carcinomas, bronchial carcinomas, and colon carcinomas. It may

XX also be used to cure cocaine addiction.

XX Sequence 413 AA;

XX Query Match 66.7%; Score 38; DB 22; Length 413;

XX Best Local Similarity 70.0%; Pred. No. 30;

XX Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

XX OY 1 SSVDSVFKP 10

XX Db 200 sldysatdp 209

Fri Jun 14 10:36:41 2002

us-09-710-339-2_copy_150_160.rag

Page 9

Search completed: June 13, 2002, 08:53:26
Job time: 276 sec

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:49:40 ; Search time 73.31 Seconds
(without alignments)
652.742 Million cell updates/sec

Title: US-09-710-339-2
Perfect score: 2684
Sequence: 1 MWAMWSLFLYGLQVAPALA.....LPRVLYPTKLAGSKICSS 498

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2684	100.0	499 1 ALASI	alpha-amyase (EC
2	2675	99.7	499 2 JS0663	alpha-amyase (EC
3	2674	99.6	499 1 ALAS3	alpha-amyase (EC
4	2674	99.6	499 2 B48305	alpha-amyase (EC
5	2673	99.6	499 2 JN0588	alpha-amyase (EC
6	2662	99.2	498 2 A48305	alpha-amyase (EC
7	2656	99.0	499 2 JT0466	alpha-amyase (EC
8	2503.5	93.3	478 2 JK0201	alpha-amyase (EC
9	1780	66.3	484 1 A35282	alpha-amyase (EC
10	1629	60.7	624 1 JC4510	alpha-amyase (EC
11	1527	56.9	507 2 S33921	alpha-amyase (EC
12	1449	54.0	512 2 S06115	alpha-amyase (EC
13	1445.5	53.9	512 2 S23355	alpha-amyase (EC
14	1397.5	52.1	484 1 ALBRYF	alpha-amyase (EC
15	1264	47.1	631 2 S72270	alpha-amyase (EC
16	1015.5	37.8	513 2 T38770	alpha-amyase (EC
17	963.5	35.9	625 2 T41603	alpha-amyase a pr
18	920.5	34.3	491 2 T38448	probable alpha-amy
19	899	33.5	581 2 S62505	probable alpha-amy
20	898.5	33.5	564 2 T41503	alpha-amyase - fi
21	898	33.5	478 2 T40860	probable alpha-amy
22	737	27.5	482 2 S31478	beta-amyase (EC 3
23	670	25.0	1196 2 A29130	beta-amyase (EC 3
24	653.5	24.3	774 2 T39539	beta-amyase homo
25	481.5	17.9	642 2 A11827	cyclomaltoedextrin
26	456.5	17.0	713 1 ALBSG1	cyclomaltoedextrin
27	454.5	16.9	712 1 ALBSG3	cyclomaltoedextrin
28	444.5	16.6	713 1 ALBSG7	cyclomaltoedextrin
29	440.5	16.4	528 1 ALBSK	alpha-amyase (EC

30	433	16.1	710 2 S63598	cyclomaltoedextrin
31	428.5	16.0	655 1 ALBSG	cyclomaltoedextrin
32	428	15.9	703 1 ALBSX1	cyclomaltoedextrin
33	424.5	15.8	919 2 S28179	alpha-amyase (EC
34	418	15.6	713 2 A58800	cyclomaltoedextrin
35	418	15.6	714 1 ALBSGR	cyclomaltoedextrin
36	414.5	15.4	713 2 S09196	cyclomaltoedextrin
37	412	15.4	717 1 S28784	alpha-amyase (EC
38	410	15.3	483 2 G75392	glycosyl hydrolase
39	407	15.2	704 2 I39805	cyclomaltoedextrin
40	406.5	15.1	718 1 ALBSGC	cyclomaltoedextrin
41	405.5	15.1	718 1 ALBSG6	cyclomaltoedextrin
42	405	15.1	713 1 ALBSMX	cyclomaltoedextrin
43	399.5	14.9	718 1 ALBSMX	cyclomaltoedextrin
44	384.5	14.3	711 1 ALBSXF	cyclomaltoedextrin
45	376.5	14.0	1104 2 A60999	alpha-amyase (EC

ALIGNMENTS

RESULT 1
ALASI
alpha-amyase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
N:Alternate names: alpha-amyase isozyme II; glycogenase; Taka-amyase A
C:Species: Aspergillus oryzae
C:Date: 30-Sep-1991 #sequence-revision 30-Sep-1991 #text-change 15-Sep-2000
C:Accession: S04548; A33214; JS0240; A91930; A93767; A10627
R:Wise, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989
A:Title: Three alpha-amyase genes of Aspergillus oryzae exhibit identical intron-exo
A:Reference number: S04548; MUID:89237897
A:Accession: S04548
A:Molecule type: DNA
A:Residues: 1-499 <MIR>
A:Cross-references: EMBL:X12725; NID:92430; PIDN:CAA31218.1; PID:9295921
A:Genetics: AMY1
A:Accession: A33214
A:Molecule type: mRNA
A:Residues: 1-499 <MIR>
A:Cross-references: GB:X12725; NID:92430; PIDN:CAA31218.1; PID:9295921
R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
Gene 79, 107-117, 1989
A:Title: Aspergillus oryzae has two nearly identical Taka-amyase genes, each contain
A:Reference number: JS0240; MUID:89378767
A:Accession: JS0240
A:Molecule type: DNA
A:Residues: 1-499 <GEN>
A:Genetics: AMY2
A:Note: the authors refer to this as isozyme II
R:Isemura, S.; Ikenaka, T.
J. Biochem. 74, 1-10, 1973
A:Reference number: A91930; MUID:74001521
A:Accession: A91930
A:Molecule type: protein
A:Residues: 206-225 <ISE>
R:Narita, K.
Proc. Jpn. Acad. 51, 285-290, 1975
A:Reference number: A93767
A:Accession: A93767
A:Molecule type: protein
A:Residues: 434-443,446-447, 'O',449-458, 'GTVV',459-464,467-468, 'B',470, 'B',472-499 <N
R:Watanabe, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984
A:Title: Structure and possible catalytic residues of Taka-amyase A.
A:Reference number: A37454; MUID:84212370
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkunburg, J.P.; Wilkinson,
submitted to the Brookhaven Protein Data Bank, August 1992
A:Reference number: A51548; PDB:67MA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics: <AMY1>

A:Gene: amy1
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Genetics: <AMY2>
 A:Gene: amy1
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; polys
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase 1 #status experimental <AMY>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:227,251,318/Active site: Asp, Glu, Asp #status predicted

231 21 323

Query Match 100.0%; Score 2684; DB 1: Length 499;
 Best Local Similarity 100.0%; Pred. No. 1,2e-185;
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MVAWMSLFLVGLQVAAAPALATPADWRSOSIFLLTDFARFDTGTTATCTADOKKCG 60
 |||||||
 2 MVAWMSLFLVGLQVAAAPALATPADWRSOSIFLLTDFARFDTGTTATCTADOKKCG 61
 |||||||
 61 TWGGIIDLKLDYIQGMGFTAIWTPVTAQLPQTAYGDAYHGYWQDIYSLNENGTADL 120
 |||||||
 62 TWGGIIDLKLDYIQGMGFTAIWTPVTAQLPQTAYGDAYHGYWQDIYSLNENGTADL 121
 |||||||
 121 KALSSALHERGMYLWVAVANHMGYDAGSSVDYSVFRPSSQDYFHPFCFIONYEQOTQ 180
 |||||||
 122 KALSSALHERGMYLWVAVANHMGYDAGSSVDYSVFRPSSQDYFHPFCFIONYEQOTQ 181
 |||||||
 181 VEDCWIGDNTVSLPDLDTTKDYVKNEMVDWVGLSVNSIDGLRIDYVKNHOKDFWPGYN 240
 |||||||
 182 VEDCWIGDNTVSLPDLDTTKDYVKNEMVDWVGLSVNSIDGLRIDYVKNHOKDFWPGYN 241
 |||||||
 241 KAAGVYCIGEVLDGDPATYCPYQNVMDGVLPYIPYLLNAFKSTSGSMDLLYNNINTVK 300
 |||||||
 242 KAAGVYCIGEVLDGDPATYCPYQNVMDGVLPYIPYLLNAFKSTSGSMDLLYNNINTVK 301
 |||||||
 301 SDCPSTLLGTVEVENDNRFASVTNDIALAKNVAAFIILNDGIPITAYGQBOHYAGND 360
 |||||||
 302 SDCPSTLLGTVEVENDNRFASVTNDIALAKNVAAFIILNDGIPITAYGQBOHYAGND 361
 |||||||
 361 PANREATWLSGYPTDSELYKLIASANAIIRNYAISKDTGFTVYKKNPIYKDDTTIAMRGT 420
 |||||||
 362 PANREATWLSGYPTDSELYKLIASANAIIRNYAISKDTGFTVYKKNPIYKDDTTIAMRGT 421
 |||||||
 421 DGSQIVTITLSNKGASGDSYTLISLGAGYTAGOQLLEVIGCTTVYVSGDNVVPVMAAGLP 480
 |||||||
 422 DGSQIVTITLSNKGASGDSYTLISLGAGYTAGOQLLEVIGCTTVYVSGDNVVPVMAAGLP 481
 |||||||
 481 RVLYPTEKLAGSKICSS 498
 |||||||
 482 RVLYPTEKLAGSKICSS 499

RESULT 2
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.

C:Species: Aspergillus sp.
 C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
 C:Accession: J50663
 R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
 Biosci. Biotechnol. Biochem. 56, 174-179, 1992
 A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its expression
 A:Reference number: J50663; MUID:9223146
 A:Accession: J50663
 A:Molecule type: mRNA
 A:Residues: 1-499 <SH1>

C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase #status predicted <ALP>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.7%; Score 2675; DB 2: Length 499;
 Best Local Similarity 99.8%; Pred. No. 5.6e-185;
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MVAWMSLFLVGLQVAAAPALATPADWRSOSIFLLTDFARFDTGTTATCTADOKKCG 60
 |||||||
 2 MVAWMSLFLVGLQVAAAPALATPADWRSOSIFLLTDFARFDTGTTATCTADOKKCG 61
 |||||||
 61 TWGGIIDLKLDYIQGMGFTAIWTPVTAQLPQTAYGDAYHGYWQDIYSLNENGTADL 120
 |||||||
 62 TWGGIIDLKLDYIQGMGFTAIWTPVTAQLPQTAYGDAYHGYWQDIYSLNENGTADL 121
 |||||||
 121 KALSSALHERGMYLWVAVANHMGYDAGSSVDYSVFRPSSQDYFHPFCFIONYEQOTQ 180
 |||||||
 122 KALSSALHERGMYLWVAVANHMGYDAGSSVDYSVFRPSSQDYFHPFCFIONYEQOTQ 181
 |||||||
 181 VEDCWIGDNTVSLPDLDTTKDYVKNEMVDWVGLSVNSIDGLRIDYVKNHOKDFWPGYN 240
 |||||||
 182 VEDCWIGDNTVSLPDLDTTKDYVKNEMVDWVGLSVNSIDGLRIDYVKNHOKDFWPGYN 241
 |||||||
 241 KAAGVYCIGEVLDGDPATYCPYQNVMDGVLPYIPYLLNAFKSTSGSMDLLYNNINTVK 300
 |||||||
 242 KAAGVYCIGEVLDGDPATYCPYQNVMDGVLPYIPYLLNAFKSTSGSMDLLYNNINTVK 301
 |||||||
 301 SDCPSTLLGTVEVENDNRFASVTNDIALAKNVAAFIILNDGIPITAYGQBOHYAGND 360
 |||||||
 302 SDCPSTLLGTVEVENDNRFASVTNDIALAKNVAAFIILNDGIPITAYGQBOHYAGND 361
 |||||||
 361 PANREATWLSGYPTDSELYKLIASANAIIRNYAISKDTGFTVYKKNPIYKDDTTIAMRGT 420
 |||||||
 362 PANREATWLSGYPTDSELYKLIASANAIIRNYAISKDTGFTVYKKNPIYKDDTTIAMRGT 421
 |||||||
 421 DGSQIVTITLSNKGASGDSYTLISLGAGYTAGOQLLEVIGCTTVYVSGDNVVPVMAAGLP 480
 |||||||
 422 DGSQIVTITLSNKGASGDSYTLISLGAGYTAGOQLLEVIGCTTVYVSGDNVVPVMAAGLP 481
 |||||||
 481 RVLYPTEKLAGSKICSS 498
 |||||||
 482 RVLYPTEKLAGSKICSS 499

RESULT 3

alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme 1; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713
 R:Wiersel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
 A:Reference number: S04549; MUID:89237897
 A:Accession: S04549
 A:Molecule type: DNA
 A:Residues: 1-499 <NR>
 A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 A:Accession: A33215
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-499 <W12>
 A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 R:Genes: M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989
A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
A:Reference number: J50240; MUID:99378767
A:Accession: A44713
A:Molecule type: DNA
A:Residues: 1-499 <GEN>
A:Note: the authors refer to this as Isoczyme I
R:Matsumura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984
A:Title: Structure and possible catalytic residues of Taka-amylase A.
A:Reference number: A37454; MUID:84212370
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Gene: amy3; Amy1
A:Introns: 56/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; metal
F:1-21/Domain: signal sequence #status predicted <SIG>
F:2-499/Product: alpha-amylase #status experimental <MAT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 99.6%; Score 2674; DB 1; Length 499;
Best Local Similarity 99.6%; Pred. No. 6,6e-185;
Matches 496; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MYAMWSLFLYGLQVAAPALATPADWRSQSYFLITDRFARTDGTATCNTADQKCGG 60
DB 2 MYAMWSLFLYGLQVAAPALATPADWRSQSYFLITDRFARTDGTATCNTADQKCGG 61
QY 61 TWGGIIDLKLDYIOGKGFALWITPVTALPOTTAAGDAHGWOODIYSLNENYGTADL 120
DB 62 TWGGIIDLKLDYIOGKGFALWITPVTALPOTTAAGDAHGWOODIYSLNENYGTADL 121
QY 121 KALSALHERGMYLWVNVANHMGTGAGSSVDYSVKPFSSODYFHPFCFIONYEDQTO 180
DB 122 KALSALHERGMYLWVNVANHMGTGAGSSVDYSVKPFSSODYFHPFCFIONYEDQTO 181
QY 181 VEDCMLGDNVTSLPDLDTTKDVKNEMVDWGSLSVNSYSDGLRIDTVKHVKQDFWPGYN 240
DB 182 VEDCMLGDNVTSLPDLDTTKDVKNEMVDWGSLSVNSYSDGLRIDTVKHVKQDFWPGYN 241
QY 241 KAAGYVCIGEVLDGDPATTCPCYQNVMDGVNLNPIYPLNLFKSTSGSMDLYNMINTVK 300
DB 242 KAAGYVCIGEVLDGDPATTCPCYQNVMDGVNLNPIYPLNLFKSTSGSMDLYNMINTVK 301
QY 301 SDCPSTLLGTFVENHNDPRFASVTNDIALAKNVAALFIINDGIPITYAGOEQHYAGND 360
DB 302 SDCPSTLLGTFVENHNDPRFASVTNDIALAKNVAALFIINDGIPITYAGOEQHYAGND 361
QY 361 PANREATWLSGYPDSELYKLIAANAI RNVAISKDGFYTKNMPYIKDDTTIPMRKGT 420
DB 362 PANREATWLSGYPDSELYKLIAANAI RNVAISKDGFYTKNMPYIKDDTTIPMRKGT 421
QY 421 DGSQIVTLLSNKGASGDSYTLISGAGYTAGOOLTEYIGCTTIVVGS DGNVPVPMAGGLP 480
DB 422 DGSQIVTLLSNKGASGDSYTLISGAGYTAGOOLTEYIGCTTIVVGS DGNVPVPMAGGLP 481
QY 481 RVLYPTEKLAGSKICSS 498
DB 482 RVLYPTEKLAGSKICSS 499

RESULT 4
BA8305

alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: B48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from As
A:Reference number: A48305; MUID:90254827
A:Accession: B48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 99.6%; Score 2674; DB 2; Length 499;
Best Local Similarity 99.6%; Pred. No. 6,6e-185;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MYAMWSLFLYGLQVAAPALATPADWRSQSYFLITDRFARTDGTATCNTADQKCGG 60
DB 2 MYAMWSLFLYGLQVAAPALATPADWRSQSYFLITDRFARTDGTATCNTADQKCGG 61
QY 61 TWGGIIDLKLDYIOGKGFALWITPVTALPOTTAAGDAHGWOODIYSLNENYGTADL 120
DB 62 TWGGIIDLKLDYIOGKGFALWITPVTALPOTTAAGDAHGWOODIYSLNENYGTADL 121
QY 121 KALSALHERGMYLWVNVANHMGTGAGSSVDYSVKPFSSODYFHPFCFIONYEDQTO 180
DB 122 KALSALHERGMYLWVNVANHMGTGAGSSVDYSVKPFSSODYFHPFCFIONYEDQTO 181
QY 181 VEDCMLGDNVTSLPDLDTTKDVKNEMVDWGSLSVNSYSDGLRIDTVKHVKQDFWPGYN 240
DB 182 VEDCMLGDNVTSLPDLDTTKDVKNEMVDWGSLSVNSYSDGLRIDTVKHVKQDFWPGYN 241
QY 241 KAAGYVCIGEVLDGDPATTCPCYQNVMDGVNLNPIYPLNLFKSTSGSMDLYNMINTVK 300
DB 242 KAAGYVCIGEVLDGDPATTCPCYQNVMDGVNLNPIYPLNLFKSTSGSMDLYNMINTVK 301
QY 301 SDCPSTLLGTFVENHNDPRFASVTNDIALAKNVAALFIINDGIPITYAGOEQHYAGND 360
DB 302 SDCPSTLLGTFVENHNDPRFASVTNDIALAKNVAALFIINDGIPITYAGOEQHYAGND 361
QY 361 PANREATWLSGYPDSELYKLIAANAI RNVAISKDGFYTKNMPYIKDDTTIPMRKGT 420
DB 362 PANREATWLSGYPDSELYKLIAANAI RNVAISKDGFYTKNMPYIKDDTTIPMRKGT 421
QY 421 DGSQIVTLLSNKGASGDSYTLISGAGYTAGOOLTEYIGCTTIVVGS DGNVPVPMAGGLP 480
DB 422 DGSQIVTLLSNKGASGDSYTLISGAGYTAGOOLTEYIGCTTIVVGS DGNVPVPMAGGLP 481
QY 481 RVLYPTEKLAGSKICSS 498
DB 482 RVLYPTEKLAGSKICSS 499

RESULT 5
JN0588
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
C:Accession: JN0588
R:Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kiritla, N.; Tsuboi, A.; Uda, S.
Gene 84, 319-327, 1989
A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
A:Reference number: JN0588; MUID:90128276
A:Accession: JN0588

A:Molecule type: mRNA
 A:Residues: 1-499 <TSD>
 C:Comment: The alpha amylases are encoded by multigene family.
 C:Genetics:
 A:Gene: Taa-GI
 A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <Sig>
 F:22-499/Product: alpha-amylase #status predicted <AMT>
 F:194-321/Domain: alpha-amylase core homology <AMT>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.6%; Score 2673; DB 2; Length 499;
 Best Local Similarity 99.6%; Pred. No. 7.7e-185;
 Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

1 MYAMWSLFLYGLQVAAPALATPADMRSOSIYFLITDRFARTDSSTATCNTADOKYCG 60
2 MYAMWSLFLYGLQVAAPALATPADMRSOSIYFLITDRFARTDSSTATCNTADOKYCG 61
3
61 TWGGIIDKLDYIQGMGFALWTPVTAQLPQTAYGDAHYGWODIYSLNENGTADDL 120
62 TWGGIIDKLDYIQGMGFALWTPVTAQLPQTAYGDAHYGWODIYSLNENGTADDL 121
63
121 KALSSALHERGMYLWVDVANHMGYDAGSSVDYVFKPFSSQDYFHPFCFIQNYEDQTQ 180
122 KALSSALHERGMYLWVDVANHMGYDAGSSVDYVFKPFSSQDYFHPFCFIQNYEDQTQ 181
123
181 VEDCWLGDNTVSLPDLDTTKDVKNKNEWYDWGSLVSNYSIDGLRIDYVKHYOKDFWPGYN 240
182 VEDCWLGDNTVSLPDLDTTKDVKNKNEWYDWGSLVSNYSIDGLRIDYVKHYOKDFWPGYN 241
183
241 KAAGYCTGEVLDGDPATYCPYQNMVGLNYPITYPLNAFKSTSGSMDDLXNMINTVK 300
242 KAAGYCTGEVLDGDPATYCPYQNMVGLNYPITYPLNAFKSTSGSMDDLXNMINTVK 301
243
301 SDPCPSTLLGTFFVENHNDNRFASYNTNDIALAKNVAAFIILNDGIPITYAGOEQHYAGND 360
302 SDPCPSTLLGTFFVENHNDNRFASYNTNDIALAKNVAAFIILNDGIPITYAGOEQHYAGND 361
303
361 PANREATWLSGYPTDSELYKLASNAIRNAISMDTGVTYKMPYIKDDTTIAMRGT 420
362 PANREATWLSGYPTDSELYKLASNAIRNAISMDTGVTYKMPYIKDDTTIAMRGT 421
363
421 DGSQIVTILSNKAGSDSYTILSGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAGGLP 480
422 DGSQIVTILSNKAGSDSYTILSGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAGGLP 481
423
481 RVLYPTEKLAGSKICSSS 498
482 RVLYPTEKLAGSKICSSS 499
483

```

RESULT 6
 A48305
 alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori
 C:Species: Aspergillus awamori
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
 C:Accession: A48305
 R:Kotman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
 Curr. Genet. 17, 203-212, 1990
 A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus
 A:Reference number: A48305; MUID:90254827
 A:Accession: A48305
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-498 <KOR>
 C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylase core homology <AMT>

Query Match 99.2%; Score 2662; DB 2; Length 498;
 Best Local Similarity 99.6%; Pred. No. 4.8e-184;
 Matches 493; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

1 MYAMWSLFLYGLQVAAPALATPADMRSOSIYFLITDRFARTDSSTATCNTADOKYCG 60
2 MYAMWSLFLYGLQVAAPALATPADMRSOSIYFLITDRFARTDSSTATCNTADOKYCG 61
3
61 TWGGIIDKLDYIQGMGFALWTPVTAQLPQTAYGDAHYGWODIYSLNENGTADDL 120
62 TWGGIIDKLDYIQGMGFALWTPVTAQLPQTAYGDAHYGWODIYSLNENGTADDL 121
63
121 KALSSALHERGMYLWVDVANHMGYDAGSSVDYVFKPFSSQDYFHPFCFIQNYEDQTQ 180
122 KALSSALHERGMYLWVDVANHMGYDAGSSVDYVFKPFSSQDYFHPFCFIQNYEDQTQ 181
123
181 VEDCWLGDNTVSLPDLDTTKDVKNKNEWYDWGSLVSNYSIDGLRIDYVKHYOKDFWPGYN 240
182 VEDCWLGDNTVSLPDLDTTKDVKNKNEWYDWGSLVSNYSIDGLRIDYVKHYOKDFWPGYN 241
183
241 KAAGYCTGEVLDGDPATYCPYQNMVGLNYPITYPLNAFKSTSGSMDDLXNMINTVK 300
242 KAAGYCTGEVLDGDPATYCPYQNMVGLNYPITYPLNAFKSTSGSMDDLXNMINTVK 301
243
301 SDPCPSTLLGTFFVENHNDNRFASYNTNDIALAKNVAAFIILNDGIPITYAGOEQHYAGND 360
302 SDPCPSTLLGTFFVENHNDNRFASYNTNDIALAKNVAAFIILNDGIPITYAGOEQHYAGND 361
303
361 PANREATWLSGYPTDSELYKLASNAIRNAISMDTGVTYKMPYIKDDTTIAMRGT 420
362 PANREATWLSGYPTDSELYKLASNAIRNAISMDTGVTYKMPYIKDDTTIAMRGT 421
363
421 DGSQIVTILSNKAGSDSYTILSGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAGGLP 480
422 DGSQIVTILSNKAGSDSYTILSGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAGGLP 481
423
481 RVLYPTEKLAGSKIC 495
482 RVLYPTEKLAGSKIC 496
483

```

RESULT 7
 JT0466
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 N:Alternate names: glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
 C:Accession: JT0466
 R:Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
 Agric. Biol. Chem. 53, 593-599, 1989
 A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus
 A:Reference number: JT0466
 A:Accession: JT0466
 A:Molecule type: DNA
 A:Residues: 1-499 <TAD>
 C:Comment: See also PIR:J0201 and PIR:J0240.
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics:
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <Sig>
 F:22-499/Product: alpha-amylase #status predicted <AMT>

F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 99.0%; Score 2656; DB 2; Length 499;
 Best Local Similarity 99.4%; Pred. No. 1,3e-183;
 Matches 495; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 MVAWMSLFLYGLQVAPALAAATPADMSOSITFLITDFARTDGTATCNTADOKYCGG 60
   |||||||
DB 2 MVAWMSLFLYGLQVAPALAAATPADMSOSITFLITDFARTDGTATCNTADOKYCGG 61
QY 61 TWOGIIDLKLDYIQGMGFTAWITPTAOLPOTAYGDAYHGYWQODIYSLNENYGTADDL 120
   |||||||
DB 62 TWOGIIDLKLDYIQGMGFTAWITPTAOLPOTAYGDAYHGYWQODIYSLNENYGTADDL 121
QY 121 KALSALHERGMYLWVDVYANMGYDAGSSVDYVFPFPSSODYFHPFCFIONYEDQTO 180
   |||||||
DB 122 KALSALHERGMYLWVDVYANMGYDAGSSVDYVFPFPSSODYFHPFCFIONYEDQTO 181
QY 181 VEDCWLGDNTVSLPDLITTKDYVKNEMWDVGSIVSNYSIDGLRIDTVKHYQKDFWPGYN 240
   |||||||
DB 182 VECWMLGDNTVSLPDLITTKDYVKNEMWDVGSIVSNYSIDGLRIDTVKHYQKDFWPGYN 241
QY 241 KAAGVVCIGEVLDGDPAYTCPTQNMWDGVLNPIYPIPLNPFKSTSGSMDDLNMINTVK 300
   |||||||
DB 242 KAAGVVCIGEVLDGDPAYTCPTQNMWDGVLNPIYPIPLNPFKSTSGSMDDLNMINTVK 301
QY 301 SDCEPDSLTLGTFEVENHONRFPASYNDIALAKNVAFTLLDNGIPIIYVAGOECHYAGCND 360
   |||||||
DB 302 SDCEPDSLTLGTFEVENHONRFPASYNDIALAKNVAFTLLDNGIPIIYVAGOECHYAGCND 361
QY 361 PANREATWISGYPTEDESEKILIASANAIIRNVAISKDGFVYKMKPIYKDDTTIAMRGT 420
   |||||||
DB 362 PANREATWISGYPTEDESEKILIASANAIIRNVAISKDGFVYKMKPIYKDDTTIAMRGT 421
QY 421 DGSQIVTILSNKASGDSYTLISGAGYTAGOQLTEVIGCTTVTGSDGNVPVMAGGLP 480
   |||||||
DB 422 DGSQIVTILSNKASGDSYTLISGAGYTAGOQLTEVIGCTTVTGSDGNVPVMAGGLP 481
QY 481 RVLYPTEKLAGSKICSS 498
   |||||||
DB 482 RVLYPTEKLAGSKICSS 499
   |||||||

```

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
 C:Accession: JK0201
 R: Toda, H.; Kondo, K.; Narita, K.
 Proc. Jpn. Acad. 58B, 208-212, 1982
 A: Title: The complete amino acid sequence of Taka-amylase A.
 A: Reference number: JK0201
 A: Accession: JK0201
 A: Molecule type: protein
 A: Residues: 1-478 <TOD>
 C: Comment: One atom of calcium per molecule is essential for the activity.
 C: Comment: This enzyme is a glycoprotein.
 C: Comment: See also PIR:J00466 and PIR:J0240.
 C: Function:
 A: Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A: Pathway: glycogen/starch degradation
 C: Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C: Keywords: glycoprotein; glycosidase; hydrolysis; polysaccharide degradation
 F:173-300/domain: alpha-amylase core homology <AMY>
 F:197/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 93.3%; Score 2503.5; DB 2; Length 478;
 Best Local Similarity 97.9%; Pred. No. 1,2e-172;
 Matches 468; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

```

QY 21 ATPADMSOSITFLITDFARTDGTATCNTADOKYCGGTWOGIIDLKLDYIQGMGFTAI 80
   |||||||
DB 1 ATPADMSOSITFLITDFARTDGTATCNTADOKYCGGTWOGIIDLKLDYIQGMGFTAI 60
QY 81 WITPVTALPQTTAYGDAYHGYWQODIYSLNENYGTADDLKALSALHERGMYLWVDVVA 140
   |||||||
DB 61 WITPVTALPQTTAYGDAYHGYWQODIYSLNENYGTADDLKALSALHERGMYLWVDVVA 120
QY 141 NHMGIDGAGSSVDYVFKPPSSODYFHPFCFIONYEDQTOVEDCWLGNITYSLPDLITK 200
   |||||||
DB 121 NHMGIDGAGSSVDYVFKPPSSODYFHPFCFIONYEDQTOVEDCWLGNITYSLPDLITK 180
QY 201 DVKNEMWDVGSIVSNYSIDGLRIDTVKHYQKDFWPGYNKAAGYVCIGEVLDGDPAYTC 260
   |||||||
DB 181 DVKNEMWDVGSIVSNYSIDGLRIDTVKHYQKDFWPGYNKAAGYVCIGEVLDGDPAYTC 240
QY 261 PYQNMWDGVLNPIYPIPLNPFKSTSGSMDDLNMINTVKSCEPDSLTLGTFEVENHNDPR 320
   |||||||
DB 241 PYQNMWDGVLNPIYPIPLNPFKSTSGSMDDLNMINTVKSCEPDSLTLGTFEVENHNDPR 300
QY 321 FASTYNDIALAKNVAFTLLDNGIPIIYVAGOECHYAGCNDPANREATWISGYPTESELYK 380
   |||||||
DB 301 FASTYNDIALAKNVAFTLLDNGIPIIYVAGOECHYAGCNDPANREATWISGYPTESELYK 360
QY 361 LIASANAIIRNVAISKDGFVYKMKPIYKDDTTIAMRGTDSQIVTILSNKASGDSYT 440
   |||||||
DB 361 LIASANAIIRNVAISKDGFVYKMKPIYKDDTTIAMRGTDSQIVTILSNKASGDSYT 419
QY 441 LSLGAGYTAGOQLTEVIGCTTVTGSDGNVPVMAGGLPRLVYPTERKLAGSKICSS 498
   |||||||
DB 420 LSLGAGYTAGOQLTEVIGCTTVTGSDGNVPVMAGGLPRLVYPTERKLAGSKICSS 477

```

RESULT 9
 A35282
 alpha-amylase (EC 3.2.1.1) - Aspergillus niger
 C:Species: Aspergillus niger
 C:Date: 10-Sep-1989 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 R: Boel, E.; Brady, L.; Brzozowski, A.M.; Derewenda, Z.; Dodson, G.G.; Jensen, V.J.; P
 Biochemistry 29, 6244-6249, 1990
 A: Title: Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-angstrom
 A: Reference number: A35282; PMID:91002514
 A: Accession: A35282
 A: Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A: Molecule type: mRNA
 A: Residues: 1-484 <BOE>
 C: Function:
 A: Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A: Pathway: glycogen/starch degradation
 C: Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C: Keywords: glycoprotein; glycosidase; hydrolysis; polysaccharide degradation
 F:173-300/domain: alpha-amylase core homology <AMY>

Query Match 66.3%; Score 1780; DB 1; Length 484;
 Best Local Similarity 67.2%; Pred. No. 1,6e-120;
 Matches 319; Conservative 63; Mismatches 93; Indels 0; Gaps 0;

```

QY 24 ADMRSOSITFLITDFARTDGTATCNTADOKYCGGTWOGIIDLKLDYIQGMGFTAIWIT 83
   |||||||
DB 4 ADMRSOSITFLITDFARTDGTATCNTADOKYCGGTWOGIIDLKLDYIQGMGFTAIWIS 63
QY 84 PVTALPQTTAYGDAYHGYWQODIYSLNENYGTADDLKALSALHERGMYLWVDVVA 143
   |||||||
DB 64 PVTALPQTTAYGDAYHGYWQODIYSLNENYGTADDLKALSALHERGMYLWVDVVA 123
QY 144 GTDAGSSVDYVFKPPSSODYFHPFCFIONYEDQTOVEDCWLGNITYSLPDLITTKDYV 203
   |||||||

```

Db 124 GYAGNGNDVSVFDPDSSSYFHPYCLITDMNLTMVQDCWEGPTIYSLPDLNTTETAV 183
 QY 204 KKNEMDWMGSLVNSISIDLRIDTFVKHVKQDFWPGYNKAGVYICGEVLDGDPATTCPYQ 263
 Db 184 RITWIDMADVLSNYSVDLRIDSVLEVPDPFPGYQEAAGVYCGVEVNDGNPALDCPYQ 243
 QY 264 NVMGVLNTPPIYPLINAFKSTSGMDLYNNINIVKSPCDPSTLLGTFFVENHNDPRAS 323
 Db 244 KVLGVLNTPYQWLLYAFESSGSSINLYNMKISVAPSCSDPTLLGNTEHNDNPRAS 303
 QY 324 YNMDILAKNVAFFILNDGIPITVYAGEOHYAGNDNPRATWLSGPTSELYKILIA 383
 Db 304 YTSIDYQANKNVSYIFLSGIPITVYAGEOHYSGGKVPINRATWLSGDTSAELTYMTA 363
 QY 384 SANAINVAISKDTGFVYKKNPIYKDDTTIAMRKGTDSQITVILSNKASGDSYTLSL 443
 Db 364 TTAIRKILIASDSAVITYANDAFYDTSNTIAMRKGTSGSVITVLSNKGSSGSSYTLTL 423
 QY 444 SGAGYTAGOOLTEVIGCTTVYSGDGNVPVPMAGGLPRVLYPTEKLAGSKICSS 498
 Db 424 SSGSYTSGTKLIEAYTCTSVYDSSGDIIPVMASGLPRVLLPRAVVDSSSLCGGS 478

RESULT 10

JC4510
 alpha-amylase (EC 3.2.1.1) precursor - yeast (Lipomyces kononenkoae)
 N:Alternate names: LKAI protein; raw starch-degrading amylase
 N:Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)
 C:Species: Lipomyces kononenkoae
 C>Date: 07-Feb-1996 #sequence_revision 23-Aug-1996 #text_change 18-Jun-1999
 C:Accession: J04510; PC4116
 R:Steyn, A.J.C.; Marmur, J.; Pretorius, I.S.
 Gene 166, 65-71, 1995
 A:Title: Cloning, sequence analysis and expression in yeasts of a cDNA containing a Lipo
 A:Reference number: J04510; MUID:96105202
 A:Accession: J04510
 A:Molecule type: mRNA
 A:Residues: 1-624 <STP>
 A:Cross-references: GB:U03076; NID:q1173536; PIDN:AAC49622.1; PID:q1173537
 A:Experimental source: Strain IGC4052B
 A:Accession: PC4116
 A:Molecule type: protein
 A:Residues: 29-44 <ST2>
 A:Experimental source: IGC4052B
 C:Genetics:
 A:Gene: LKAI
 C:Function:
 A:Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Lipomyces alpha-amylase; alpha-amylase core homology; glucosylase starch
 C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
 C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
 F:28/Domain: signal sequence #status predicted <SIC>
 F:29-624/Product: alpha-amylase #status predicted <MAT>
 F:48-141/Domain: glucosylase starch-binding domain homology <SBD>
 F:320-447/Domain: alpha-amylase core homology <AMY>
 F:177-185,297-311,387-430,587-622/Disulfide bonds: #status predicted
 F:304,344/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:337,377,444/Active site: His, Glu, Asp #status predicted

Query Match 60.7%; Score 1629; DB 1; Length 624;
 Best Local Similarity 60.6%; Pred. No. 1.8e-109;
 Matches 292; Conservative 77; Mismatches 113; Indels 0; Gaps 0;

QY 15 AAPALATPADMRSSQSYFLITDFARPDTGTTATCNAQKYGSGTWOGITDKLDYIOG 74
 Db 142 SASVPTGTAAWKRSTIQVYTDRAFRDGSITTSCTDVTDVYGGSTGRTINMLDYIOG 201
 QY 75 MGTATWITPTVAOLPOTTAAGDAVHGWMODIYSLNENYGTADLALASALBERGMYL 134
 Db 202 MGTATWISPTVEINTPDGTGAYAHGWMADIFALNTNREGADDLALATAEHLNRGMYL 261
 QY 135 MVDVANHMGYDAGSSVDYSVFRPSSQDYFHPCFIQNTEDQTOVEDCWLGNVTSLP 194

Db 262 MVDIVVHHPAFSGNHADVISEYFPYSSQDYFHSFCMIDTYSNQNTVEECMLDSDSVLY 321
 QY 195 DLDITTKDVAKNEMWVGLSVNSISIDLRIDTFVKHVKQDFWPGYNKAGVYICGEVLDG 254
 Db 322 DVNTQDLTFKSEYOSVSKOLIANYSIDGLRIDTFVKHVKQDFWPGYNKAGVYICGEVLDG 381
 QY 255 DPATCEYQANMGOVLNPIYPLINAFKSTSGMDLYNNINIVKSPCDPSTLLGTFFVE 314
 Db 382 DPSTCEYQENLDGVLPYIPYPLINAFKSTSGMDLYNNINIVKSPCDPSTLLGTFFVE 441
 QY 315 NHDNPRASYNTDIALAKNVAFFILNDGIPITVYAGEOHYAGNDNPRATWLSGPT 374
 Db 442 NQDNPRPYSYSDSESLKILIAFTLISDGIPIYVYAGEOGLNGNDPNYREALMPGYSY 501
 QY 375 DSELYKILIASANAINVAISKDTGFVYKKNPIYKDDTTIAMRKGTDSQITVILSNK 434
 Db 502 TSTFEYVYASLNOIRNHAIVYIDTYLQWVYISDSTTTIAMRKGTGNOITVLSMGS 561
 QY 435 SGDSYTLISGAGYTAGOOLTEVIGCTTVYSGDGNVPVPMAGGLPRVLYPTEKLAGSK 494
 Db 562 SSGSYTTLTSLNLTGTTASVYVEILLTCAVYVDLSGNLAVPMAGGLPRVYVESOLVSG 621
 QY 495 CS 496
 Db 622 CS 623

RESULT 11

S33921
 alpha-amylase (EC 3.2.1.1) SWA2 precursor - yeast (Schwanniomyces occidentalis)
 N:Alternate names: alpha-1,4 glucanohydrolase
 C:Species: Schwanniomyces occidentalis
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S33921
 R:Claros, M.G.; Abarea, D.; Fernandez-Lobato, M.; Jimenez, A.
 Curr. Genet. 24, 75-83, 1993
 A:Title: Molecular structure of the SWA2 gene encoding an AM1-related alpha-amylase
 A:Reference number: S33921; MUID:93365041
 A:Accession: S33921
 A:Molecule type: DNA
 A:Residues: 1-507 <CTA>
 A:Cross-references: EMBL:X73497; NID:93396561; PIDN:CAA51912.1; PID:93396562
 C:Genetics:
 A:Gene: SWA2
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-507/Product: alpha-amylase #status predicted <MAT>
 F:205-332/Domain: alpha-amylase core homology <AMY>
 F:134,222/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 56.9%; Score 1527; DB 2; Length 507;
 Best Local Similarity 57.7%; Pred. No. 3e-102;
 Matches 286; Conservative 74; Mismatches 130; Indels 6; Gaps 3;

QY 1 MYAAMWELPIFXGQVAPALAAATPADMRSSQSYFLITDFARPDTGTTATCNAQKYGCG 60
 Db 17 LVASKRPIFLSKRDAGSAA--WRSSTIQLVTDRAFRDGSITSATCNTGMDVYCGG 72
 QY 61 TWOGITDKLDYIOGAGFTATWITPTVAOLPOTTAAGDAVHGWMODIYSLNENYGTADL 120
 Db 73 TPGGITDKLDYIOGAGFTATWISPVVEQIPDDTGAYAHGWMADIFALNTNREGADDL 132
 QY 121 KALSSALBERGMYLWVDVANHMGYDAGSSVDYSVFRPSSQDYFHPCFIQNTEDQTO 180
 Db 133 KNLSENELHRKNMKLWVDIYTNHYNAGAGSSVAYSNVFNPOOSYFHDCLLTNDDQTN 192
 QY 181 VEDCWLGNVTSLPDLITTKDVAKNEMWVGLSVNSISIDLRIDTFVKHVKQDFWPGYN 240

```

Db 193 VEDWEDBNVTYSLPDLTETEDSDVSSIFNLVAVELVSNYSIDLRLIDSAKHVEDSEFSPSQ 252
OY 241 KAAGVYCGIEVLDDPAPYTCYQONVMDGYLVNPIYVPLLNARKSTSGSMDLNNYNIVK 300
Db 253 SAAGVYILGVEYDDPAPYTCYQONVMSGYVNPFLYPLMLRFQGSNSVDELNAMISLE 312
OY 301 SDPDSSTLLGTEVENHNPFRFASYNIDIALAKNVAAPILINDGIPPIYAGOEQHYAGND 360
Db 313 SDCKDITLLGFENHNPRLPSYSDSALIKNAIFNLMDSGIDIPPIYAGOEQHYAGND 372
OY 361 PANREATWLSGYPDSLEYKLIASANARYAISKDTGVTYKKNPIYKDDTTIMRGT 420
Db 373 PNNREALMLSGYSTNSGYTKLISVNOIRNOAIYKDSKTTTWSDVYLASGHVIALORGA 432
OY 421 DGSQIVTILSKGASGDSYTLISGAGYTAGOQLREVIGCTVTYVSGDNVPVPMAGSLP 480
OY 433 DDQRTVSVFNLLGSSG-SQYVTFSS-TRYSQGEKVVDVLICQTSYANSDBTLTVSISGAP 490
OY 481 RVLYPTEKLASKTICS 496
Db 491 RYAPASTLIANSIGICN 506

```

RESULT 12

S06115

alpha-amylose (EC 3.2.1.1) AMY1 precursor - yeast (Schwanniomycetes occidentalis)

C:Species: Schwanniomycetes occidentalis

C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 22-Jun-1999

C:Accession: S06115

R:Strasser, A.W.M.; Selk, R.; Dohmen, R.J.; Niermann, T.; Bielefeld, M.; Seeboth, P.; Tu

Eur. J. Biochem. 184, 699-706, 1989

A:Title: Analysis of the alpha-amylose gene of Schwanniomycetes occidentalis and the secre

A:Reference number: S06115; MIMD:90032659

A:Accession: S06115

A:Molecule type: DNA

A:Residues: 1-512 <STR>

A:Cross-references: EMBL:X16040; NID:g4882; PIDN:CAA34162.1; PID:g4883

C:Genetics:

A:Gene: AMY1

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycoegen/starch degradation

C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

C:1-25/Domain: signal sequence #status predicted <SIS>

C:26-512/Product: alpha-amylose #status predicted <AMT>

C:209-336/Domain: alpha-amylose core homology <AMT>

Query Match 54.0%; Score 1449; DB 2; Length 512;

Best Local Similarity 56.4%; Pred. No. 1.3e-96;

Matches 269; Conservative 74; Mismatches 132; Indels 2; Gaps 2;

```

OY 21 ATPAD-WRSOSIYFLLRFRARTGTTATCNTADQKYGCGTMOGIITDKLDYIOGMGFPA 79
Db 36 SSSADKMKDQSIYQIVTRFARSDGTTADCLVSDRKYCGSGTKGIITDKLDYIOGMGFPA 95
OY 80 IWTPTVTAQLPQTTAYGDAYHGYWQODIYSLNENYGTADLKLASSALHERGMYLWVAV 139
Db 96 IWTPTVTAQLPQTTAYGDAYHGYWQODIYSLNENYGTADLKLASSALHERGMYLWVAV 155
OY 140 ANHMGYDAGSSVDYVFKPPSSODYFHPRCFIONYEDTOYVDCWLGNTYSLPDLSTE 199
Db 156 YNHIAWNGDSSVDYVFKPPSSODYFHPRCFIONYEDTOYVDCWLGNTYSLPDLSTE 215
OY 200 KDQVKNEMYDMWGLSVNYSIDGLRIDTVKHKQDFWPGYNKAGVYCGIEVLDDPAPY 259
Db 216 DNEVIGVFOQVWSDVFNQYVSIIDGLRIDSAKHVDTASLTKFEDASGYVNLGEVYQGPPT 275
OY 260 CPYQVWMDGYLVNPIYVPLLNARKSTSGSMDLNNYNIVKSCDPSSTLLGTFVENHNP 319
Db 276 CPYQVWMDGYLVNPIYVPLLNARKSTSGSMDLNNYNIVKSCDPSSTLLGTFVENHNP 335

```

RESULT 13

S23355

alpha-amylose (EC 3.2.1.1) precursor - yeast (Schwanniomycetes occidentalis)

C:Species: Schwanniomycetes occidentalis

C:Date: 22-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999

C:Accession: S23355

R:Mu, F.M.; Wang, T.T.; Hsu, W.H.

FEMS Microbiol. Lett. 82, 313-318, 1991

A:Title: The nucleotide sequence of Schwanniomycetes occidentalis alpha-amylose gene.

A:Reference number: S23355

A:Accession: S23355

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-512 <MUR>

A:Cross-references: EMBL:X62079; NID:g4880; PIDN:CAA43995.1; PID:g4881

A:Note: the authors translated the codon AGA for residue 21 as Pro, CTT for residue 6

C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology

C:Keywords: glycoprotein; glycosidase; hydrolase

F:209-336/Domain: alpha-amylose core homology <AMT>

Query Match 53.9%; Score 1445.5; DB 2; Length 512;

Best Local Similarity 56.5%; Pred. No. 2.2e-96;

Matches 266; Conservative 74; Mismatches 130; Indels 1; Gaps 1;

```

OY 26 WRSOSIYFLLRFRARTGTTATCNTADQKYGCGTMOGIITDKLDYIOGMGFPAITTPV 85
Db 42 WKQOSIYQIVTRFARSDGTTADCLVSDRKYCGSGTKGIITDKLDYIOGMGFPAITSPV 101
OY 86 TAQLPQTTAYGDAYHGYWQODIYSLNENYGTADLKLASSALHERGMYLWVAVANHMY 145
Db 102 VEQIPNTAYGDAYHGYWQODIYSLNENYGTADLKLASSALHERGMYLWVAVNYHAM 161
OY 146 DAGSSVDYVFKPPSSODYFHPRCFIONYEDTOYVDCWLGNTYSLPDLTKDYKN 205
Db 162 NGDSSVDYVFKPPSSODYFHPRCFIONYEDTOYVDCWLGNTYSLPDLTKDYKN 221
OY 206 EWDVWGLSVNYSIDGLRIDTVKHKQDFWPGYNKAGVYCGIEVLDDPAPYTCYQNV 265
Db 222 VEQVWSDVFNQYVSIIDGLRIDSAKHVDTASLTKFEDASGYVNLGEVYQGPPT 281
OY 266 MGVVLYNPIYVPLLNARKSTSGSMDLNNYNIVKSCDPSSTLLGTFVENHNPFRAS 325
Db 282 MKGVVLYNPIYVPLLNARKSTSGSMDLNNYNIVKSCDPSSTLLGTFVENHNPFRAS 341
OY 326 NDIALAKNVAAPILINDGIPPIYAGOEQHYAGNDPANREATWLSGYPDSLEYKLIA 385
Db 342 SDPSLILKNMAAFILINDGIPPIYAGOEQHYAGNDPANREATWLSGYPDSLEYKLIA 401
OY 386 NAIKNVAAPILINDGIPPIYAGOEQHYAGNDPANREATWLSGYPDSLEYKLIA 445
Db 402 NAIKNVAAPILINDGIPPIYAGOEQHYAGNDPANREATWLSGYPDSLEYKLIA 460
OY 446 AGYTAGOQLREVIGCTVTYVSGDNVPVPMAGSLPRLVYPTTEKLASKTICS 496
Db 461 TQYSSGDKVYIILISQNSVYAGDPFSLVSIISGMPQYAYRASSVLSGSGICN 511

```

RESULT 14

ALBYAF
alpha-amylose (EC 3.2.1.1) precursor - yeast (*Saccharomyopsis fibuligera*)
C:Species: *Saccharomyopsis fibuligera*
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 18-Jun-1999
C:Accession: S00064
R:Itoh, T.; Yamashita, I.; Fukui, S.
FEBS Lett. 219, 339-342, 1987
A:Title: Nucleotide sequence of the alpha-amylose gene (ALP1) in the yeast *Saccharomyopsis fibuligera*
A:Reference number: S00064; MUID:87276512
A:Accession: S00064
A:Molecule type: DNA
A:Residues: 1-494 <IT0>
A:Cross-references: EMBL:X05791; NID:94847; PIDN:CAA29233.1; PID:94848
A:Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in having 368-Thr instead of 368-Ile.
C:Genetics:
A:Gene: ALP1
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
C:Keywords: calcium; extracellular protein; glycoprotein; glycosidase; hydrolase; metallopeptidase
C:EC:3.2.1.1-6/Domains: signal sequence #status predicted <SIG>
C:87-494/Product: alpha-amylose #status predicted <MAT>
C:200-337/Domains: alpha-amylose core homology <AMY>
E:57-65,177-191,267-310,462-493/Disulfide bonds: #status predicted
E:148,189,202,237/Binding site: calcium (Asn, Gln, Asp, His) #status predicted
E:224/Binding site: carbohydrate (Asn) (covalent) #status predicted
E:233,257,324/Active site: Asp, Gln, Asp #status predicted

```

Query March 1997.51%; Score 1397.5; DB 1; Length 494;
Best Local Similarity 52.5%; Pred. 6.1e-93;
Matches 264; Conservative 83; Mismatches 129; Indels 27; Gaps 6;

QY 12 LQVAAPALAA-----TPAD-WRSQSIFLLTDFRFRATDGSSTACNTA 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 M0ISKALALSLAALVYAQPVTLFFKRETNADKRQSQIQTITDFRFRATDGSASCNTE 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 54 DOKYGGTWMOGIIDLVDYIOGKGFATMTPTTAQLPQTAYGADYHGYWODIYSLMEN 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 DRLYGGSSFGQIKRLKDKGFTAIMISPYEENIPDNTAAGYAHGYWKMNIKIMEN 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 YGTADDLKALSALHERGMYTLMDVYVANHMGYDGAAGSSVDYSEVKPESQDIFFHPCTIQ 173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 FGTADDLKSLMOELHDMRLMVMYDVTNHYGSDGSDGSDIDSEYRPFNDOKFHFNYCLIS 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 NYEEDTQYEDDCGLGNFTVSLPDLDTTKDYVKNEWDMWGSILVNSITGLRIIDTYKHQVK 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 NYDDQAQVQSCHEGSSVALPDLRTEDSDVASVFMSWKDPFGANSISDGLRTDSAKHNQDQ 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 DFWPGYKKAAGYCYIGCEVLIDGDPATTCYQNVMDCVLNTPIYPLLNFAKSTGSSMDLY 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 GFPPDVFVASGYSYGEVFGDPATTCYQNTIPVSNSTPIYPTTRFFKFTDSSSESLT 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 NMINTVAKSDCPSTLTGTFVENHNDNRPFASYNIDALAKNAFAFTLLNDGPIIYIAGOEQ 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 QMISSVASSCSDPVLTFNFEVNHONERFASMSDQSLISNALFAVLLGDIPIVYYGQD 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 HYAGNDPANRRATMLSGYPTDSELYKLIALSANALIRNAISKDQGFVYIKKMPYIKDDTT 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 GLTSGSDPNRRREALWSGKNKESDYKLIAKANAARNAAYVODSSYATVSQLSVITSNDHV 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 414 IAMRGKTGSGSOIVTLLSNKKGASGDSYTLISAGATYGAQQLFEVIGCTTVYGSGDNV 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 IATKRGs----VVSYPNNLGSAGSS--DVIISMVYSSGDELEVLTGCTSVSSGSD--LQY 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 474 PMAGGLPRVLYPTEKLAGSKTCS 496
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 474 SIQGGQPIQFVPAK--VYASDICS 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15
S72270

alpha-amylase EC 3.2.1.1) precursor - *Cryptococcus* sp. (strain CS2)
C:Species: *Cryptococcus* sp.
A:Variety: strain CS2
C:Date: 23-Apr-1998 #sequence-revision 08-May-1998 #text-change 20-Jun-2000
C:Accession: S72270
R:Refuji, H.; Chino, M.; Kato, M.; Iimura, Y.
Biochem. J. 318, 989-996, 1996
A:Title: Raw-starch-digesting and thermostable alpha-amylase from the yeast *Cryptococcus*
A:Reference number: S72270; MUID:96433120
A:Accession: S72270
A:Molecule type: DNA
A:Residues: 1-631 <IEF>
A:Cross-references: EMBL:D83540; NID:g1595852; PID:BAAL2010.1; PID:g1595853
A:Experimental source: strain S-2
C:Genetics:
A:Gene: amy-CS2
A:Introns: 289/2, 326/2
C:Function:
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-30/Domains: signal sequence #status predicted <SIG>
F:21-63/Product: alpha-amylase #status predicted <MAT>
F:206-335/Domains: alpha-amylase core homology <AMY>

Query Match	Similarity	49.28%	Pred. No. 3.6e-83	Matches 251	Conservative 84	Mismatches 133	Indels 22	Gaps
9	LYGLVAAPALAA--TPADMRSSOSIYELLTDRFARTDSTTATGCTADQ--KYCGGTWOG	64						
Db	LAGLALLSLTSLVAGLSPAEMRSSOSIYQVYVTRDFALDNGNSPSCGSESLNLKCNCFAG	66						
Qy	IIDLDVYIQMGFTAIWTPYTAQLP-----QTTAYGAYAGWQOQDILSNENYGTAD	119						
Qy	65							
Db	IIDLDVYIQMGFTAIWTPYTAQLP-----QTTAYGAYAGWQOQDILSNENYGTAD	119						
Qy	120							
Db	LKALSSALHBERGMYLAVDVVANHMGV-----DGAGSSVDYSVFKPPRSSODYFHPFCFI	172						
Qy	127							
Db	LTTDSLNLHSGRYLWAVDVVNMHAYLCCTDGCSPGNSVNGSTPTPNSSEYHPCEI	186						
Qy	173							
Db	QNEEDQVQVEECMLGDNVTSYSLPDLDTTKVYVKNENYMDWGSIVSNYSIDGLRIDTYVAHQ	232						
Qy	187							
Db	DYNNRRISILDCMEGDEIVPLDRLKREDDVGSIFNSMSINSLQITNYNDGLRIDSILQSG	245						
Qy	223							
Db	KDFMPGTNKA--GVYCIQEVLDGDPATYTCPTYNV--MDGLVLPITYIPLLNKFKSTSSMD	290						
Qy	246							
Db	SFFPPGPNQAAGNAYMYGEVFNSSPEYVCPYQOAAQMPGLVLPMEFYTTNAPFQSSGS	305						
Qy	291							
Db	DLYNMIVTASDCPDGSLTLOTPEVENDNDRFASYNIDTALAKNNAALFIINDGIPIIYAG	350						
Qy	306							
Db	QLAGGISAMQSDSDPTLLSGFLFENQDNRRFPQSJSLDTRAQNMALFAPMLQDDGIPITYG	365						
Qy	351							
Db	QEOHYAGAGNDPANREATWLS--GYPTDSELYKILASANAIRYATSKDGTGYTYNNMPTIK	409						
Qy	366							
Db	QEOHSSGSGVPLNKEALMTISGVDTSIPLEYEMITTVNOLRLTALIKONGEFTYRIQVPT	425						
Qy	410							
Db	DDTTIARKKTDTSQIYTIISNKGAGSDYSYLSLGA--GYTAGQOLEVYIGCTTVVGS	467						
Qy	426							
Db	DSNHIVTRKKNSGIQTIVYTNVNGSAGSSSTLISSESEFGQASBPVMDVLSCTLYHTGT	485						
Qy	468							
Db	DGNVNVPMAGGLPRVLPPTTEKLAGSKTCS	497						
Qy	486							
Db	DGSLSTMTGTGLPRVEYNATATALESSLCT	515						

Search completed: June 13, 2002, 08:55:45
Job time: 369 sec

Fri Jun 14 10:36:40 2002

us-09-710-339-2.rpt

Page 9

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:53:35 ; Search time 125.81 Seconds
(Without alignments)
684.774 Million cell updates/sec

Title: US-09-710-339-2
Perfect score: 2684
Sequence: 1 MVAWMSLFYGLQVAPALAA.....LPRVLPTPEKLAGSKICSSS 498

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2674	99.6	499	3	Q96TH4	Q96TH4 aspergillus
2	1826	68.0	490	3	Q9UV07	Q9UV07 emericella
3	1810	67.4	640	3	Q13236	Q13236 aspergillus
4	1725.5	64.3	623	3	Q9UV09	Q9UV09 emericella
5	1629	60.7	624	3	Q01117	Q01117 lipomyces k
6	1527	56.9	507	3	Q08806	Q08806 debaryomyces
7	1264	47.1	631	3	Q92394	Q92394 cryptococcus
8	963.5	35.9	625	3	Q74922	Q74922 schizosacch
9	920.5	34.3	491	3	Q13956	Q13956 schizosacch
10	898.5	33.5	564	3	Q9Y759	Q9Y759 schizosacch
11	737	27.5	482	2	Q60051	Q60051 thermococci
12	653.5	24.3	774	3	Q42918	Q42918 schizosacch
13	493.5	18.4	323	3	Q96WR6	Q96WR6 schizosacch
14	453	16.9	1798	2	Q9K211	Q9K211 streptomyces
15	447.5	16.7	713	2	Q9F5M3	Q9F5M3 bacillus ci
16	428	15.9	725	2	Q59239	Q59239 bacillus sp

17	418.5	15.6	739	1	Q9UWN2	Q9UWN2 thermococcus
18	418	15.6	714	2	Q52766	Q52766 paenibacill
19	410	15.3	483	16	Q9RUB8	Q9RUB8 deinococcus
20	408	15.2	704	2	Q82984	Q82984 bacillus sp
21	404.5	15.1	1358	2	P70983	P70983 bacillus sp
22	403.5	15.0	692	2	Q30565	Q30565 bacillus br
23	399	14.9	711	16	Q992B3	Q992B3 streptococ
24	393.5	14.7	1021	2	Q9K5L6	Q9K5L6 actinoplan
25	392.5	14.6	711	2	Q9ZAO0	Q9ZAO0 bacillus st
26	378	14.1	588	2	Q65007	Q65007 thermus sp.
27	376.5	14.0	1104	2	Q06812	Q06812 micrococcus
28	373.5	13.9	580	16	Q9A609	Q9A609 caulobacter
29	369	13.7	588	2	Q9A1V2	Q9A1V2 bacillus st
30	368	13.7	588	2	Q93R22	Q93R22 bacillus st
31	367.5	13.7	724	2	Q9K5L5	Q9K5L5 actinoplan
32	363	13.5	690	16	Q9K186	Q9K186 vibrio chol
33	343.5	12.8	677	2	Q9RHR1	Q9RHR1 klebsiella
34	336.5	12.5	578	16	Q9K856	Q9K856 bacillus ha
35	333	12.4	1684	2	Q03658	Q03658 unidentified
36	332	12.4	589	2	Q9R9H8	Q9R9H8 bacillus su
37	332	12.4	589	16	Q06988	Q06988 bacillus su
38	329.5	12.3	958	16	Q9KFR4	Q9KFR4 bacillus ha
39	328.5	12.2	567	16	Q99ZB2	Q99ZB2 streptococ
40	324	12.1	558	2	Q59226	Q59226 bacillus sp
41	320	11.9	587	2	Q82982	Q82982 bacillus sp
42	320	11.9	591	16	Q929P3	Q929P3 listeria in
43	315.5	11.8	590	2	Q45490	Q45490 bacillus st
44	307.5	11.5	526	2	Q60102	Q60102 xanthomonas
45	305	11.4	955	2	Q45574	Q45574 bacillus sp

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	499 AA.
Q96TH4	Q96TH4	01-DEC-2001 (TREMBLrel. 19, Created)		
AC	Q96TH4	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).			
GN	AMYA OR AMYI.			
OS	Aspergillus oryzae, and			
OS	Aspergillus flavus.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.			
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=5062, 5059;			
RP	SEQUENCE FROM N.A.			
RP	SPECIES=A.oryzae; STRAIN=RIH40;			
RX	MEDLINE=20289310; PubMed=10830498;			
RA	Gomi K., Akeno T., Minehoki T., Ozeki K., Kumagai C., Okazaki N.,			
RA	Iimura Y.;			
RT	"Molecular cloning and characterization of a transcriptional activator			
RT	gene, amyR, involved in the amylolytic gene expression in Aspergillus			
RT	oryzae.";			
RL	Biosci. Biotechnol. Biochem. 64:816-827(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	SPECIES=A.flavus; STRAIN=86-10D;			
RC	Fakhoury A.M., Molosnuk C.P.;			
RT	"Amyl, the alpha-amylase gene of Aspergillus flavus: Involvement in			
RT	afatoxin biosynthesis in maize kernels.";			
RL	Phytopathology 83:908-914(1993).			
DR	EMBL; AB021876; BAA95703.1; -			
DR	EMBL; AF139925; AAF14264.1; -			
SO	SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;			

Query Match 99.6%; Score 2674; DB 3; Length 499;
Best Local Similarity 99.6%; Pred. No. 2.3e-174;
Matches 496; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MYAMMSLFYGLQVAPALAAATPADMRSSQSYFLTDRFARTDGSSTATCMTADQKCYG 60
   |||||
DB 2 MYAMMSLFYGLQVAPALAAATPADMRSSQSYFLTDRFARTDGSSTATCMTADQKCYG 61
QY 61 TWOGIIDKLDYIQGMGFTAIWITPVTAOLPQTAYGDAYHGYOODIYSLNENYGTADL 120
   |||||
DB 62 TWOGIIDKLDYIQGMGFTAIWITPVTAOLPQTAYGDAYHGYOODIYSLNENYGTADL 121
QY 121 KALSSALHERGMVLMVDVYANHMGYDAGSSVDYVFKPFSSODYFHPCLQNTEDQTO 180
   |||||
DB 122 KALSSALHERGMVLMVDVYANHMGYDAGSSVDYVFKPFSSODYFHPCLQNTEDQTO 181
QY 181 VEDCMLGDMTVSLPDLDTTKDVKNEMVWVSGSLVSNYSIDGLRIDTVKHYQKDPWPGYN 240
   |||||
DB 182 VEDCMLGDMTVSLPDLDTTKDVKNEMVWVSGSLVSNYSIDGLRIDTVKHYQKDPWPGYN 241
QY 241 KAAGVYCIGEVLDGDPAYTCYQVNMWDGVLNPIYYPLLNAPKSTSGSMDDLXNMIINTVK 300
   |||||
DB 242 KAAGVYCIGEVLDGDPAYTCYQVNMWDGVLNPIYYPLLNAPKSTSGSMDDLXNMIINTVK 301
QY 301 SDCCPDTLLGTVEVNDNDRFASNTYNDIALAKNVAFFILLNDGIPITAYAGQOHYAGGND 360
   |||||
DB 302 SDCCPDTLLGTVEVNDNDRFASNTYNDIALAKNVAFFILLNDGIPITAYAGQOHYAGGND 361
QY 361 PANREATWLSGYPDTSELYKLIAANAIKNVAISKDTGFTYKKNPIYKDDTTIAMRKGT 420
   |||||
DB 362 PANREATWLSGYPDTSELYKLIAANAIKNVAISKDTGFTYKKNPIYKDDTTIAMRKGT 421
QY 421 DGSQIVTILSNKGASGDSYTLISLGAGYTAGOOLTEVIGCTTVTVGSDGNVPMAGLIP 480
   |||||
DB 422 DGSQIVTILSNKGASGDSYTLISLGAGYTAGOOLTEVIGCTTVTVGSDGNVPMAGLIP 481
QY 481 RVLYPTEKLAGSKICSS 498
   |||||
DB 482 RVLYPTEKLAGSKICSS 499

RESULT 2
QY 09UV07 PRELIMINARY; PRT; 490 AA.
AC 09UV07;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ALPHA-AMYLASE AMYA.
GN AMYA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eucristales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RA Boase N.A., Murphy R.M., Kelly J.M.;
RT "An amylase cluster in Aspergillus nidulans."
RT Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF208225; AAF17103.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha_amylase.
DR Pfam: PF00128; alpha-amylase; 1.
DR Pfam: PF02806; alpha-amylase_C; 1.
DR PRINTS: PR00110; ALPHAAMYLASE.
SQ SEQUENCE 490 AA; 54249 MW; A891CACEAEB5305 CRC64;

```

Query Match 68.0%; Score 1826; DB 3; Length 490;
 Best Local Similarity 69.6%; Pred. No. 1.5e-116;
 Matches 337; Conservative 53; Mismatches 92; Indels 2; Gaps 2;

```

QY 12 LOVAALALATPADMRSSQSYFLTDRFARTDGSSTATCMTADQKCGGTWGIIDKLDY 71
   |||||
DB 7 LSFATVLAATPAEMRSQSYFLTDRFARTDGSSTATCMTADQKCGGTWGIIDKLDY 65

```

```

QY 72 IQGMGFTAIWITPVTAOLPQTAYGDAYHGYOODIYSLNENYGTADL KALSSALHERG 131
   |||||
DB 66 IQGMGFTAIWITPVTAOLPQTAYGDAYHGYOODIYALNPBHGTDJRALSDALHDBG 124
QY 132 MYLWADVANHMGYDAGSSVDYVFKPFSSODYFHPFCIQYEDQTOYEDCWLQDNVY 191
   |||||
DB 125 MYLWADVANHMGYDAGSSVDYVFKPFSSODYFHPFCIQYEDQTOYEDCWLQDNVY 184
QY 192 SLPDLDTTKDVKNEMVWVSGSLVSNYSIDGLRIDTVKHYQKDPWPGYNKAAGVYCIGEV 251
   |||||
DB 185 SLPDLDTTKDVKNEMVWVSGSLVSNYSIDGLRIDTVKHYQKDPWPGYNKAAGVYCIGEV 244
QY 252 LDGDPAYTCYQVNMWDGVLNPIYYPLLNAPKSTSGSMDDLXNMIINTVKSDCCPDTLLGT 311
   |||||
DB 245 LDGDPAYTCYQVNMWDGVLNPIYYPLLNAPKSTSGSMDDLXNMIINTVKSDCCPDTLLGT 304
QY 312 FVENHNDNDRFASNTYNDIALAKNVAFFILLNDGIPITAYAGQOHYAGGNDPANREATWLSG 371
   |||||
DB 305 FVENHNDNDRFASNTYNDIALAKNVAFFILLNDGIPITAYAGQOHYAGGNDPANREATWLSG 364
QY 372 YPTDSELYKLIAANAIKNVAISKDTGFTYKKNPIYKDDTTIAMRKGT DGSQIVTILSN 431
   |||||
DB 365 YPTDSELYKLIAANAIKNVAISKDTGFTYKKNPIYKDDTTIAMRKGT DGSQIVTILSN 424
QY 432 KGASGDSYTLISLGAGYTAGOOLTEVIGCTTVTVGSDGNVPMAGLIPRVLYPTEKLAG 491
   |||||
DB 425 AGADAGSSYTVSVPNTGTTAGAAVTEIYTCEDITVSGSGEVSMESGLPRVLYPTEKLAG 484
QY 492 SKIC 495
   |||||
DB 485 SKIC 488

```

```

RESULT 3
QY 013296 PRELIMINARY; PRT; 640 AA.
AC 013296;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ACID-STABLE ALPHA-AMYLASE.
OS Aspergillus kawachii.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eucristales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=40384;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaneko A.;
RT Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kaneko A., Sudo S., Sakamoto Y., Tamura G., Ishikawa T., Ohba T.;
RT "Molecular cloning and determination of the nucleotide sequence of a
RT gene encoding an acid-stable alpha-amylase from Aspergillus-
RT kawachii."
RT J. Ferment. Bioeng. 81:292-298(1996).
DR EMBL: AB008370; BAA2293.1; -.
DR HSSP: P56271; 2AAA.
DR InterPro: IPR000461; Alpha_amylase.
DR InterPro: IPR002044; CBD_4.
DR Pfam: PF00128; alpha-amylase; 1.
DR Pfam: PF00686; CBD_4; 1.
DR ProDom: PD001568; CBD_4; 1.
SQ SEQUENCE 640 AA; 65507 MW; 676BF9D0236720DD CRC64;

```

Query Match 67.4%; Score 1810; DB 3; Length 640;
 Best Local Similarity 67.6%; Pred. No. 2.6e-115;
 Matches 325; Conservative 66; Mismatches 90; Indels 0; Gaps 0;

```

QY 18 ALAATPADMRSSQSYFLTDRFARTDGSSTATCMTADQKCGGTWGIIDKLDYIQGMG 77
   |||||
DB 19 ALGSAEMRWQSYFLTDRFARTDGSSTATCMTADQKCGGTWGIIDKLDYIQGMG 78

```


DR	InterPro: IPR000461; Alpha.amylase.
DR	Pfam: PF00128; alpha-amylase; 1.
KM	Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Calcium; Glycoprotein.
KM	Signal.
FT	CHAIN
FT	ACT_SITE
FT	ACT_SITE
FT	ACT_SITE
FT	ACT_SITE
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	DISULFID
FT	CARBOHYD
FT	CARBOHYD
FT	SEQUENCE
28	ALPHA-AMYLASE.
29	BY SIMILARITY.
30	BY SIMILARITY.
31	BY SIMILARITY.
32	BY SIMILARITY.
33	BY SIMILARITY.
34	BY SIMILARITY.
35	BY SIMILARITY.
36	BY SIMILARITY.
37	BY SIMILARITY.
38	BY SIMILARITY.
39	BY SIMILARITY.
40	BY SIMILARITY.
41	BY SIMILARITY.
42	BY SIMILARITY.
43	BY SIMILARITY.
44	BY SIMILARITY.
45	BY SIMILARITY.
46	BY SIMILARITY.
47	BY SIMILARITY.
48	BY SIMILARITY.
49	BY SIMILARITY.
50	BY SIMILARITY.
51	BY SIMILARITY.
52	BY SIMILARITY.
53	BY SIMILARITY.
54	BY SIMILARITY.
55	BY SIMILARITY.
56	BY SIMILARITY.
57	BY SIMILARITY.
58	BY SIMILARITY.
59	BY SIMILARITY.
60	BY SIMILARITY.
61	BY SIMILARITY.
62	BY SIMILARITY.
63	BY SIMILARITY.
64	BY SIMILARITY.
65	BY SIMILARITY.
66	BY SIMILARITY.
67	BY SIMILARITY.
68	BY SIMILARITY.
69	BY SIMILARITY.
70	BY SIMILARITY.
71	BY SIMILARITY.
72	BY SIMILARITY.
73	BY SIMILARITY.
74	BY SIMILARITY.
75	BY SIMILARITY.
76	BY SIMILARITY.
77	BY SIMILARITY.
78	BY SIMILARITY.
79	BY SIMILARITY.
80	BY SIMILARITY.
81	BY SIMILARITY.
82	BY SIMILARITY.
83	BY SIMILARITY.
84	BY SIMILARITY.
85	BY SIMILARITY.
86	BY SIMILARITY.
87	BY SIMILARITY.
88	BY SIMILARITY.
89	BY SIMILARITY.
90	BY SIMILARITY.
91	BY SIMILARITY.
92	BY SIMILARITY.
93	BY SIMILARITY.
94	BY SIMILARITY.
95	BY SIMILARITY.
96	BY SIMILARITY.
97	BY SIMILARITY.
98	BY SIMILARITY.
99	BY SIMILARITY.
100	BY SIMILARITY.
101	BY SIMILARITY.
102	BY SIMILARITY.
103	BY SIMILARITY.
104	BY SIMILARITY.
105	BY SIMILARITY.
106	BY SIMILARITY.
107	BY SIMILARITY.
108	BY SIMILARITY.
109	BY SIMILARITY.
110	BY SIMILARITY.
111	BY SIMILARITY.
112	BY SIMILARITY.
113	BY SIMILARITY.
114	BY SIMILARITY.
115	BY SIMILARITY.
116	BY SIMILARITY.
117	BY SIMILARITY.
118	BY SIMILARITY.
119	BY SIMILARITY.
120	BY SIMILARITY.
121	BY SIMILARITY.
122	BY SIMILARITY.
123	BY SIMILARITY.
124	BY SIMILARITY.
125	BY SIMILARITY.
126	BY SIMILARITY.
127	BY SIMILARITY.
128	BY SIMILARITY.
129	BY SIMILARITY.
130	BY SIMILARITY.
131	BY SIMILARITY.
132	BY SIMILARITY.
133	BY SIMILARITY.
134	BY SIMILARITY.
135	BY SIMILARITY.
136	BY SIMILARITY.
137	BY SIMILARITY.
138	BY SIMILARITY.
139	BY SIMILARITY.
140	BY SIMILARITY.
141	BY SIMILARITY.
142	BY SIMILARITY.
143	BY SIMILARITY.
144	BY SIMILARITY.
145	BY SIMILARITY.
146	BY SIMILARITY.
147	BY SIMILARITY.
148	BY SIMILARITY.
149	BY SIMILARITY.
150	BY SIMILARITY.
151	BY SIMILARITY.
152	BY SIMILARITY.
153	BY SIMILARITY.
154	BY SIMILARITY.
155	BY SIMILARITY.
156	BY SIMILARITY.
157	BY SIMILARITY.
158	BY SIMILARITY.
159	BY SIMILARITY.
160	BY SIMILARITY.
161	BY SIMILARITY.
162	BY SIMILARITY.
163	BY SIMILARITY.
164	BY SIMILARITY.
165	BY SIMILARITY.
166	BY SIMILARITY.
167	BY SIMILARITY.
168	BY SIMILARITY.
169	BY SIMILARITY.
170	BY SIMILARITY.
171	BY SIMILARITY.
172	BY SIMILARITY.
173	BY SIMILARITY.
174	BY SIMILARITY.
175	BY SIMILARITY.
176	BY SIMILARITY.
177	BY SIMILARITY.
178	BY SIMILARITY.
179	BY SIMILARITY.
180	BY SIMILARITY.
181	BY SIMILARITY.
182	BY SIMILARITY.
183	BY SIMILARITY.
184	BY SIMILARITY.
185	BY SIMILARITY.

RP	[1]	SEQUENCE FROM N.A.
RC	STRAIN-ATCC	26077;
RX	MEDLINE-93365041;	Pubmed-8358835.
RA	Claros M.G.,	Abarca D., Fernandez-Lobato M., Jimenez A.;
RT	"Molecular structure of the Sma2 gene encoding an Amy1-related alpha-	
RI	amylase from Schwannomyces occidentalis."	
RL	Curr. Genet.	24:75-83(1993).
CC	-I- CATALYTIC ACTIVITY:	ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC	LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.	
CC	-I- COFACTOR:	BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC	-I- ENZYME REGULATION:	ALPHA-AMYLASE EXPRESSION UNDERLIES CATABOLITE
CC	REPRESSION BY GLUCOSE.	
CC	-I- SIMILARITY:	BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.	
DR	EMBL:	X73497; CAAS1912.1; '-.
DR	HSSP:	P10529; 7TAA.
DR	InterPro:	IPR000461; Alpha.amylase.
KM	Pfam:	PF00128; alpha-amylase; 1.
KM	HydroLase:	Glycosidase; Carbohydrate metabolism; Calcium;
KM	Glycoprotein.	
FT	ACT_SITE	238 238 BY SIMILARITY.
FT	ACT_SITE	242 242 BY SIMILARITY.
FT	ACT_SITE	329 329 BY SIMILARITY.
FT	DISULFID	62 70 BY SIMILARITY.
FT	DISULFID	182 196 BY SIMILARITY.
FT	DISULFID	272 315 BY SIMILARITY.
FT	DISULFID	470 505 BY SIMILARITY.
FT	CYSORDER	229 229 N-LINKED (GLCNAC. . .) (PROBABLE).
SQ	SEQUENCE	507 AA; 55966 MW; 3A562B95BD8AND53 CRC64;

Query Match	56.9%	Score 1527	DB 3	Length 507
Best Local Similarity	57.7%	Pred. No. 3.8e-96		
Matches 286	Conservative	74	Mismatches 130	Indels 6
				Gaps 3
QY	1 MYAMMSLEFYGLQVAPALAAPADMRQSOISIFELLTDFRFAFDGSGTATACNTADKRYCGG 60			
Db	17 LVASKRIFELSKDAGSSAAA----WRSSIIQVLYTDRAFRIDGSGTASACNIGDRYCGG 72			
QY	61 TWOGIIDLKLDYIQGMGFTAIWITPYTAQLPPTTANGDAYHGYWOODIYSLNENYGTADDL 120			
Db	73 TFGGIDTKLDYIQGMGFTAIWISPVVEQIPDDTGYGYAAHGYWMDIYAINSNFTADDL 132			
QY	121 KALSSALHRRGWYIAMWYVANAEMHGVCAGSSVDYSEYKFEPSDDGYHPHPCFIQNYEDDQO 180			
Db	133 KNLNSLELHRNNKMLAWDIYTNHRVAMGAGSSVAYSNSNPFNMQOSTPHDYCLITNDDQTN 192			
QY	181 VEDCWIGMDVTSLPDLDDITVKDYKKNEMWYMWGLSYNSYSIDGLRIJDYKHYOKDFWPNGYN 240			
Db	193 VEDCWIGMDVTSLPDLRTEDSDVSSIFNLMAVELVSNYSIDGLRIDSAKHDESPFSPQ 252			
QY	241 KAAGYVCIGEVLDGDPAYTCPYQONVMDGYLNPYYPLINAERKSTSGSMDLLYMNATYK 300			
Db	253 SAAGVYLLEGEVYDGDPAAYCPYQNYMSGVTNPPLYXPMLEFPGTSGNSVDELNAMATSLSE 312			
QY	301 SDCPDSLTCTFPEENEDNDRFASYTEMDILAKVNAFILLNGIPIITYGEOEHYAGND 360			
Db	313 SDCKDTLLGNLEFNHDDQPLPBSDYSALIKALINFLMSDIPILITYGQEDGEGYSSD 372			
QY	361 PANREATWLSGYPTDESELYKLASANAANNYAISKDTGEVYTKNMPDIYKDDTTIAMRKGT 420			
Db	373 PNNRKELMLTSGYSTSNQYKLLISVANGINNOALIKRQSKTYYTVDVLYASGHVIALORGA 432			
QY	421 DGSQIVTILSNKAGSDSYTLLSGAGYTAGOOLTEVIGCTTYTVGSDGNAVPPYMAIGLP 480			
Db	433 DDQRIYSVEFNLLGSSG-SCYTFPS-TKYSGGEKVYDVLTCQTSYANSDSSTLITVISIGGAP 490			
QY	481 RUYLPTEKLAGSKTCS 496			
Db	491 RIYAPASLIANSICN 506			

```

RESULT 7
Q92394 092394 PRELIMINARY; PRT; 631 AA.
AC 092394;
DT 01-FEB-1997 (TReMBrel. 02, Created)
DT 01-FEB-1997 (TReMBrel. 02, last sequence update)
DT 01-JUN-2001 (TReMBrel. 17, last annotation update)
DE ALPHA-AMYLASE PRECURSOR.
GN AMY-CS2.
OS Cryptococcus sp. S-2.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
OC mitosporic Hymenomycetes; Cryptococcus.
OC NCBI_TaxId=87049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-2;
RX MEDLINE=96431120; PubMed=8836148;
RA Iefuji H., Chino M., Kato M., Iimura Y.;
RT "Raw-starch-digesting and thermostable alpha-amylase from the yeast
RT Cryptococcus sp. S-2: purification, characterization, cloning and
RT sequencing.";
RL Biochem. J. 318:989-996(1996).
DR EMBL; D83541; BAA12011.1; -.
DR EMBL; D83540; BAA12010.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR002044; CBD_4.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF00686; CBD_4; 1.
DR ProDom; PD001568; CBD_4; 1.
DR Signal.
KW SIGNAL.
FT CHAIN 1 20 ALPHA-AMYLASE.
SQ SEQUENCE 631 AA; 67658 MW; 819687B6E1D707E5 CRC64;

Query Match 47.1%; Score 1264; DB 3; Length 631;
Best Local Similarity 49.2%; Pred. No. 4.4e-78;
Matches 251; Conservative 84; Mismatches 153; Indels 22; Gaps 9;

QY 9 LYGQVAPALAA--TPADRSOSITFLIDRFARIDGSTATACNRADQ--KTCGGTNG 64
DB 7 LAGALLASLGVLVAGLSPEWRSOSITGVYDFRALDNGSPSCSQSELTMLTCNGTFAG 66
QY 65 IIDKLDYIQMGFTAIWTPVTAQLP-----QTAYGDAHYGWOODIYSLNENYGTAD 119
DB 67 IIDKLDYIQNGFTAIWTPVTAQLP-----QTAYGDAHYGWOODIYSLNENYGTAD 126
QY 120 LKAISALHERGATLMDVYVANHNGY-----DGAGSSVDYSEVFKPFSSQDYFHPFCF 172
DB 127 LTDSNALSHSRGMYLMDVYVANHNGYCGTDGCGGPGNSVNGSFTPFENSEYFHPFCF 186
QY 173 QNVEDOTQVECCWLGDMVTSLPDDDTTKDYVKNEMWDVWGSLSVNSIDLRIDTYKHQY 232
DB 187 -DYNNKRTSIDCWGDELIVPLVDLTEDSDVQSIFNSMISNLQTYNIDLRIDSLQDS 245
QY 233 KDFMPEGKAA-GVYCIGEVLIDGPAYTCPYQNV-MDGVINLPITYYPLLNKFKSTGSM 290
DB 246 SEFFPFGFQAAGMYMGEVFNQSPSYVCPYQAGMGVILNYPMEFYTNAFOTSSGMS 305
QY 291 DLYNNINNVKSDCPDSTLIGFEVENHNDNPRPASTNDIALAKVNAAFITINDGIPIT 350
DB 306 QIAGCISMSQSDSTTLIGSFLENQNDNPRPSTSLTRAQNALATMLQDSIPITTYG 365
QY 351 QEOHYAGGNDPANREATWLS-GYPTDSELYKLJASANAIRNVAISKDTGVTYKKNPIYK 409
DB 366 QEOHLISGSGVPLNREALMTSGDYDTSSPLYEMITVQRLTLAKONGGVTVKIQVPT 425
QY 410 DDTITAMKKGTDGSOIYTIISNKGASDSTYLSLGA--GYTAQOQILEVIGCTTYT 467
DB 426 DSNHIVTKRKGNSGYOIVGVYTNVSGASSSTLSISSETGFQASEPVMVLSCTLYHTGT 485
QY 468 DGNVPVPMAGGLPRVLYPTKLAGSKICSS 497

```

```

DB 486 DGSLSFTMGGLPRVLYNATALLAESLCTT 515

RESULT 8
Q74922 074922 PRELIMINARY; PRT; 625 AA.
AC 074922;
DT 01-NOV-1998 (TReMBrel. 08, Created)
DT 01-NOV-1998 (TReMBrel. 08, last sequence update)
DT 01-DEC-2001 (TReMBrel. 19, last annotation update)
DE ALPHA-AMYLASE.
GN SPOC757.12.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxId=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Lyne M., Rajadream M.A., Barrell B.G., Bothe G., Pohl T;
RX Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031825; CA21237.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
SQ SEQUENCE 625 AA; 68087 MW; 30A58CE6DB0C07B3 CRC64;

Query Match 35.9%; Score 963.5; DB 3; Length 625;
Best Local Similarity 40.7%; Pred. No. 1.3e-57;
Matches 201; Conservative 83; Mismatches 195; Indels 15; Gaps 9;

QY 2 VAMNSLF-LYGQVAPALAA--TPADRSOSITFLIDRFARIDGSTATACNRADQKTCGG 60
DB 6 IALPSLFLGLP---SLAKSSEWDRITLYQVITDRFA-VSDMTFPCSPDSSYCG 61
QY 61 TWOGIIRKLDYIQMGFTAIWTPVTAQLPOT--TAYGDAHYGWOODIYSLNENYGTAD 119
DB 62 TWSGIRKLDYIQMGFTAIWTPVTAQLPOT--TAYGDAHYGWOODIYSLNENYGTAD 121
QY 120 LKAISALHERGATLMDVYVANHNGYDGAGSSVDYSEVFKPFSSQDYFHPFCFQ--NYEDQ 178
DB 122 LIDLTDMHNDMMIMFDALANSMAIPGPDNITSYSLVYFNDSSYFHPYCWIDYGSNN 181
QY 179 TQVEDCWTGDMVTSLPDDDTTKDYVKNEMWDVWGSLSVNSIDLRIDTYKHQKDFM 238
DB 182 TDIEDCWTGDMVTSLPDDDTTKDYVKNEMWDVWGSLSVNSIDLRIDTYKHQKDFM 241
QY 239 YNKAAGVYCIGEVLIDGPAYTCPYQNVMDGVINLPITYYPLLNKFKSTGSM 298
DB 242 YTSAGVAFALIGEMTSYDPNVSVCVRYNLDITSPFHOGIEFAFNTGAFAFEID 301
QY 299 VKSDC--PDSTLIGFEVENHNDNPRPASTNDIALAKVNAAFITINDGIPIT 356
DB 302 FQDACEGQDMSVIGNFLENHDLPRYTSITNDTSQDIGAIFLLHNGGIPLIYYGGE 361
QY 357 GGNP-PANREATWLSGPTDSELYKLJASANAIRNVAISKDTGVTYKKNPIYK 415
DB 362 GGSOTPFERKALNMYGDTIDTANITYQTIRIALRLKQALISDSBWTDSYLDIRAH 421
QY 416 MKRGTDGSOIYTIISNKGASDSTYLSLGA--GYTAQOQILEVIGCTTYT 467
DB 422 VRKG---DVGIVYTNVSGASSSTLSISSETGFQASEPVMVLSCTLYHTGT 485
QY 476 AGGLPRVLYPTKLAGSKICSS 497
DB 477 VSGLPVYTPASL 490

RESULT 9
Q13996 013996

```

```

ID 013996 PRELIMINARY; PRT; 491 AA.
AC 013996;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
PR PUTATIVE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1).
GN SPAC27B2.01.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR EMBL; 298978; CAB11675.1; -.
DR HSSP; P10529; 77RA.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KW Glycoprotein; Signal.
FT SIGNAL 1 491 POTENTIAL.
FT CHAIN ? 491 PUTATIVE ALPHA-AMYLASE.
FT ACT_SITE 221 221 BY SIMILARITY.
FT ACT_SITE 225 225 BY SIMILARITY.
FT ACT_SITE 313 313 BY SIMILARITY.
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 491 AA; 56678 MW; 2013808D0CEB581D CRC64;

Query Match 34.3%; Score 920.5; DB 3; Length 491;
Best Local Similarity 38.7%; Pred. No. 8.2e-55;
Matches 181; Conservative 79; Mismatches 185; Indels 23; Gaps 6;

OY 26 WRSOSYFLIDRFRKDTGTTATCNTADOKYCGGTGOGIIDKIDYIOGSGFTAIWTTPY 85
DB 24 WRROCTYQILDHFDHCHTAP--STGRMTLGSTWGIIOKDYIOGSGCTAWISPI 81
OY 86 TAPDPTTAGDYGAGHYGWOQDIYSLNENYGTADDLKALSSALHBRGMYLWVVAHMGY 145
DB 82 VKNIGVGYGCAHYGWAEDTLNPHFTKODLTETLVDLHRRNMCMIDIVNHMAH 141
OY 146 DGAGSVYVYKPPSSODYPHPCFIONYEDQOVEDCWMGDMFTVSLPDLDTKDYVN 205
DB 142 AG-DSFIDISKYAPNNSSHHPRKFLNHYDIDWDCETAMIGDEVSLMDRTEDQEVHN 200
OY 206 EYDWVGLSVNSYIDGRIDTVKHVKQDFWPGYNKAGVYCIQEVLDGDPAYTCPYQNV 265
DB 201 FPGNMIRLDIQYHFHGDGRIDTAKHVKQEPYFPIAANVPAFGEVYHGPCKFIATLEY 260
OY 266 MGVLTPIYYPPLNAFSTSGSMDDL-NMINTYKSCPEPSTLSTGFVEHNDPRASY 324
DB 261 ISGANVPLTYQIENTFEPKQSMNIFOKALILEARATSMQTTILIGFTEHNDPRFLNR 320
OY 325 TNDIALAKVAVAFILINDGPIPIYAGQOHYAGNDPANRATWMLSGYPTDSELYKFIAS 384
DB 321 STDYSLCTLTLLFTFTGIPRIIFGQGOYMAAGHDPENRALMTSNVNOGNPIFFQFLKK 380
OY 385 ANAIRNVAISKDTGVYTKNMPDIYDPTTAMRKGTDSQIVTILSKGASGDSYTL--- 441
DB 381 LTRKRFIDVNSYGFTELSTNMLFVNEHYVYFR---PGVILVVSNAAGSSVDVTSAEF 436
OY 442 -----SLSGAGYTAGOOLTEVIGCTITVYGSDSGVNVPYPMAGGLPVL 483
DB 437 SITERSLEFDIVLSGSDF-----SLPTEDSSTISNMLFSPKRYL 478

```

```

RESULT 10
OY7S9
ID 09Y7S9 PRELIMINARY; PRT; 564 AA.
AC 09Y7S9;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
PR PROBABLE ALPHA-AMYLASE C63.02C PRECURSOR (EC 3.2.1.1) (1.4-ALPHA-D-
DE GLUCAN GLUCANHYDROLASE).
GN SPC63.02C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR EMBL; A1049522; CAB40006.1; -.
DR HSSP; P10529; 77RA.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KW Calcium; Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 1 21 PROBABLE ALPHA-AMYLASE C63.02C.
FT DOMAIN 537 540 POLY-SER.
FT ACT_SITE 229 229 BY SIMILARITY.
FT ACT_SITE 322 322 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 172 188 BY SIMILARITY.
FT DISULFID 263 306 BY SIMILARITY.
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 564 AA; 63205 MW; 1229PDAEC702FD0 CRC64;

Query Match 33.5%; Score 898.5; DB 3; Length 564;
Best Local Similarity 38.8%; Pred. No. 3.1e-53;
Matches 188; Conservative 85; Mismatches 190; Indels 21; Gaps 7;

OY 18 ALAAPADWRSOSYFLIDRFRKDTGTTATCNTADOKYCGGTGOGIIDKIDYIOGSGF 77
DB 19 ANAGSNAMERKRIIYQIILTRFAVDDGSTDNPCPDANOTCGGWKCIENKDIYIDMGF 78
OY 78 TALWITPYTAOLP-QTTAYGDYHGYWQODIYSLNENYGTADDLKALSSALHBRGMYLW 136
DB 79 NAIWISPIDKNIEGIDGAGYAGHYGWTYDESINHEHFGTEDDIVSLITAAHAGIWMVL 138
OY 137 DVVNAHMGYDAGGSVDYVYKPPSSODYPHPCFI--QNVEDQOVEDCWMLDGNTVSLP 194
DB 133 DSVSMALAPPLADLSSLNPNKESYFHPYCLIMDITIDNETNVMDCW-ODSGVLLA 197
OY 195 DLDITKDYVKNKMYDWVGLSVNSYIDGRIDTVKHVKQDFWPGYNKAGVYCIQEVLDG 254
DB 199 DLDVSSSVSYLSDFHFKSLISKYDQDLRIDAVKMMNYFFEPFVATGATVYSVGEVFSY 257
OY 255 DPATYCPYQNVMDGVLTNPIYPLINAFKSTSGSMDDLNMINTYKSCD--DSTLGLTF 312

```

```

Db 258 DEDYMCYSWVLPQVNTFLQDYINFSFATGACFTLLPTQVQEVNVAWSCXYDSTLMLEF 317
OY 313 VENNHNPFRASTYNDIALAKNVAFAFLINDGIPITVAGQEOHYAGNDPANREATWLSGY 372
Db 318 IENHDLRPPTYSQSQSQMGALSTVLLMDGIPSTFYGOEGFNCGEDPANPALMLTDY 377
OY 373 PDSSELYKLIASANAIRNVAISKDGFVYTKNMPYIKDDTTIARRKGTND---GSOIYV 427
Db 378 DQSNPFYVYIKTMVAFRKRVITQDQDWT-----STQGISAVDHYGQKNQDYL 428
OY 428 ILSNKGASDSTLLSGAGYTAGOOLTEVIGCTVTVGSDGNVPVPAAGLPRVLYPTE 487
Db 429 MNNNGVYTNMLTIYEVTNTYANVSDVFGHRLTIVGADKXTLTASMTNGYPLIMYDHS 487
OY 488 KLAG 491
Db 488 KMSG 491

```

```

RESULT 11
060051
ID 060051 PRELIMINARY; PRT; 482 AA.
AC 060051.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1).
GN AMYV.
OS Thermocotilomyces vulgaris.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Thermocotilomyces.
OX NCBI_TaxID=2026;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K94:
RA MEDLINE=95031040; PubMed=7944369;
RA Hofmeister B., Koenig S., Hoang V., Engel J., Mayer G., Hansen G.,
RA Hofmeister J.;
RT "The gene amyE(TV1) codes for a nonglucogenic alpha-amylase from
RT Thermocotilomyces vulgaris 94-2A in Bacillus subtilis."
RL Appl. Environ. Microbiol. 60:3381-3389(1994).
DR EMBL: X69807; CAI49465.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase: 1.
DR Pfam: PF02806; alpha-amylase-C: 1.
KM Signal: Hydrolase; Glycosidase.
FT CHAIN 1 29 POTENTIAL.
FT SIGNAL 1 29 ALPHA-AMYLASE.
SQ SEQUENCE 482 AA; 54483 MW; D8F98C370B579025 CRC64;

```

Query Match 27.5%; Score 737; DB 2; Length 482;
 Best Local Similarity 36.8%; Pred. No. 2.6e-42;
 Matches 185; Conservative 77; Mismatches 189; Indels 52; Gaps 15;

```

OY 1 MWAMSLFLYGLQVAPALPAATPADMSOSYFELTDFRARTDGTST---ATCMTAD-Q 55
Db 10 LVCILSLVLMRSTAPRTVALSPADMOEKSTIFTITDFDNGDSDNDYGGFNADKSDPR 69
OY 56 KYCGGTGOGIIDLKLDYIOGMGFTAIWTPVTAQLPQTAYGDAYHGYQODYSLENYV 115
Db 70 KYHGDFOGIILRLDYIKMGFTAIWTPVIMQ-----KSANAYHGYWTYDFYSVDHLG 124
OY 116 TADDLKALLSALHERGMLNDVYVANHMGYDAGASSVYSEKPPSSQDYHHPFCFIQNY 175
Db 125 SAAKLQELVRAAHAKGISVMDVYVANHNG-DFOBPSSFAKA---PDKADWYHHYGDITQDW 180
OY 176 EDQTOVEDCMGLDNTVSLPDLDTTKDYVKNEMVYDVGSLVNSYIDGLRITVYHYKQDF 235
Db 181 NQGMWVEN---GD-IAGLDLDLNDQNPRAVATELKNYIAHLVQTTGVDGRVUTVYAHVPEKF 236
OY 236 WFGYKKAAGVYICIGEVLDGDGPATYTCPYQNVMDGVLYNPDIYPLNLF--KSTSGSMDDLX 293

```

```

Db 237 WREDEGAANFTLGEVPHGCDPAVYGVDTYNLVDLDEPMYITTKKNFGQDQSRITADR 296
OY 294 NMATVSDCPDSTLLGTPEVNDHNPFASATYNDIALA-----KNVAAPITLLDGI 345
Db 297 AODRYRYKNPMTN---GLFIDNHVPFRLNEASCKPQASMDKMPOLKAALGFEMFTRGIP 352
OY 346 IIVAGQOHVAGNDPANREATWLSGYPTDSELYKLIASANAIRNVAISKDGTGFVYTKN 405
Db 353 ILYGTEQGSFGADPYNRD---MWFNDKHELYTAKLNYIRNHVALRKCTQAEK-- 407
OY 406 PIYKDDTTIARRKGTDSQIYTLISNKGASDSTYLLSGAGYTA-GOOLTEVIGCTVY 464
Db 408 --WVDDTFEYAFQRSYNGDEYVYMIN--SWNSQTRVPVNDIVANGQILYNNMGIDKY 462
OY 465 VGSQGNVPVPAAGLPRVLYPTE 487
Db 463 VNN-----GSITATLGPK 476

```

```

RESULT 12
042918
ID 042918 PRELIMINARY; PRT; 774 AA.
AC 042918;
DT 01-JAN-1999 (TREMBLrel. 09, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE ALPHA-AMYLASE C16A3.13 PRECURSOR (EC 3.2.1.1) (1.4-ALPHA-D-
DE GLUCAN GLUCANOHYDROLASE).
GN SPBC16A3.13.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -1- SIMILARITY: TO THE FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS
CC THE ALPHA-AMYLASE FAMILY.
DR EMBL: AL021748; CAI16864.1; -.
DR HSSP: P56271; 2AAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase: 1.
KM Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KM Calcium; Glycoprotein; Signal.
FT CHAIN 1 24 POTENTIAL.
FT SIGNAL 1 24 PUTATIVE ALPHA-AMYLASE C16A3.13.
FT ACT SITE 461 461 BY SIMILARITY.
FT ACT SITE 553 553 BY SIMILARITY.
FT DISULFID 52 60 BY SIMILARITY.
FT DISULFID 229 314 BY SIMILARITY.
FT DISULFID 495 539 BY SIMILARITY.
FT DISULFID 690 724 BY SIMILARITY.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 774 AA; 89459 MW; 92CEAAE45664269B CRC64;

```


RT the 8 Mb Streptomyces coelicolor A3(2) chromosome";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL355752; CAB90874.1; -.
 DR HSSP: P10529; 27AA.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR000461; Alpha_amyase.
 DR Pfam: PF00128; alpha-amyase; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
 SO SEQUENCE 1798 AA; 194106 MW; 43D1386BACF5CA CRC64;

Query Match 16.9%; Score 453; DB 2; Length 1798;
 Best Local Similarity 26.9%; Pred. No. 3.6e-22;
 Matches 163; Conservative 66; Mismatches 211; Indels 166; Gaps 26;

OY 19 LAATPA--DWESQSYFLTLDRFARTD-----GSTATCATADQK-----YCGGTWQGI 65
 DB 45 LAKTAARHDLTRQGFYVLDRFRANGAGNDRGITGTRLTGTYDTPDKGFYGGDLKGL 104
 OY 66 IDKLDYIOGMGFIAIWTPPYTAQLPQTTAYGDA---YHGWOODIYSLNENGTADDLKA 122
 DB 105 TERLDYIKGLGTSTIMWAPIFKNQPVQGTGKDSAGYHGWTDFQVDPHFETNKDLKN 164
 OY 123 LSSALHERGMYLWADVYANHWG--YDGAGSSVDY---SVF-----KPPSODY----- 165
 DB 165 LISKRAHAKGMVFVDVITNHTADVVDYEKSYDLSKGAFFYLTKDQGFPPDDADYADGER 224
 OY 166 -----FHPECFIQYEDQTOVEDCWLGDNTV-----S 192
 DB 225 REPVDSGSEPRPTVPAAKKNLKPMSLNDPAMYHNKGDST-----MAESATYGDENG 279
 OY 193 LPDLDTTADVYKNEWYDWGSL-----VSNSIDGLRIDTYKHVOKDWPYKNAAGY 245
 DB 260 LDDLMTPEPY-----VGMKETYQWVEDFAIDGFRIDTYKHDMEWTQMATALDA 332
 OY 246 YC-----IGEVLDGDPATYCPY--QNVMDGVLYNPITYPLNAFKSTGSMDDLY 293
 DB 333 YAAKGRDPEFMGEVYSADTSVAPYVQGRIDSTLDFP--FODAARAYASOGSARKLA 391
 OY 294 NMI-----NIVKGDGCPDSTLGFVENHNDNPRASY-----TNDIALAKN---VAAF 337
 DB 392 AFGGDYTYTTDKANAEQV---TFLGNDHMGRTGTLKDDAPBAGDAELKDRLANEL 448
 OY 338 IILNDGIPITYAGOEHOYAG--GNP-----ANREATWL-----SGYPTDS 376
 DB 449 MFLSGNPNVITYYGDGOGFTGAGGDKDARQPFASRTADYLDLDDQLGTRTHAAAYDTSA 508
 OY 377 ELIKLIASANNIR--NYAISKDGTGFYVYKKNPIYKDDTTIAMKRGTSQIYVITLSNKA 434
 DB 509 PLYROIISALAEIRKANPALA-----DGVOTERYAADGAGIYAFSRTDAK 552
 OY 435 SGGSYTSLSGAGY-----TAGOOLTEVIGC--TTTVGSDGNVPVPMAGLPRVLY 484
 DB 553 TGTETVVAFNNAAGTEKSAFAFGSAGMTFRGLYGTDAFVSGADSKYTVVPARSAVLYK 612
 OY 485 PTEKLA 490
 DB 613 AAGRLA 618

RESULT 15
 O9F5W3 PRELIMINARY; PRT; 713 AA.

AC O9F5W3;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CYCLODEXTRIN GLUCANOSYLTRANSFERASE.
 OS Bacillus circulans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OK NCBI_TaxID=1397;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=All;
 RA Rlimphanchayakit V., Tonozuka T., Sakano Y.;
 RT "Cloning of Cyclodextrin Glucanoyltransferase Gene from Bacillus
 RT circulans All.",
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF302787; AAC31622.1; -.
 DR HSSP: P05618; 1PAM.
 DR InterPro: IPR000461; Alpha_amyase.
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amyase; 1.
 DR Pfam: PF02806; alpha-amyase_C; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR PRINTS: PR00110; ALPHAMYLASE.
 DR ProDom: PD001568; CBD_4; 1.
 KW Transferase.
 SO SEQUENCE 713 AA; 78271 MW; 4FF365E1F8F93BAD CRC64;

Query Match 16.7%; Score 447.5; DB 2; Length 713;
 Best Local Similarity 27.7%; Pred. No. 2.4e-22;
 Matches 146; Conservative 77; Mismatches 215; Indels 89; Gaps 21;

OY 1 MWAMWSLFL-YGLOVAPALAA-----TPADMRQSYFLTLDRF-----ART 42
 DB 7 LTAWVTLMLSTFLSLSPVHAAPDTSVSKONFSTVYIQTIDRPSDGNPANNPTGAAF 66
 OY 43 DGSFTATCNFADQKRCGCTWOGIIDLKD--YIOGMGFIAIWTP---VTAQLPQTTAYG 96
 DB 67 DGSCNLT-----RLYCGDWOGIINKINDGYLTGGMGTAITWISQPAENITYSVINSGVHN 121
 OY 97 DAYHGYWOODIYSLNENYGTADDLKALSSALHERGMYLWADVYANHWG----- 144
 DB 122 TAYHGYWABDRKKTTPAYGTWQDFKNLIDTAHANIKVILIDFAPNHTSPASSDDPSFAEN 181
 OY 145 ---YDGAGSSVDYSVFPPSSODYFH-----PFCFION--YEDQTOVEDCWLGDNTVSLP 194
 DB 182 GRLYDNGMLGCTY-----NPTQNLFHHYGGTDFSTLNGCIYKMLYLDADLNHNNSVDY- 236
 OY 195 DLDITKDVYKNEWYDWGSLVSNSTIDGLRIDTYKHV---QKDFPQYKKAAGYCTIE 250
 DB 237 ---YLKDAIK--MWLD-----LGVDGIRVDVAKHMPFGWQKSFSTINNKRPVTFEE 284
 OY 251 VLDDGPATYCPYQNVMD---GVLNYPITYPLNAFKSTGSMDDLYNMINTVKSDDPS 306
 DB 285 MFLGVNIEISPEYHOFANESGSLDDEFRAQKARQVFRDNDNMYGLKAMLESSEVDYAGY 344
 OY 307 TLGTFVENHNDNPRASYTNDIALAKNVAFFIINDGIPITYAGOEHOYAGNDPANREA 366
 DB 345 NDQVYFINDHMERPHHSNGDRRLLEQALAFITLSRQVPAIYYSQDYMSGGNDPNRAR 404
 OY 367 TWLSGYPTDSELYKLIASANNIR--NYAISKDGTGFYVYKKNPIYKDDTTIAMKRGDSQ 424
 DB 405 --IPSFSTTTAYQYIQKLAFLKRSNPAIAGS---IQERW--INNVIYIEKKPENNVA 457
 OY 425 IVTILSNKG--ASGDSYTLISLGAGYTAGOOLTEVIGCTTVYVGSBG 469
 DB 458 VVAINRNNTPTASITGLVTSIPQGSY--NDVLGGILNGMTLVFVAGAG 502

Search completed: June 13, 2002, 09:11:06
 Job time: 1051 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:48:50 ; Search time 158.68 Seconds
(without alignments)
348.593 Million cell updates/sec

Title: US-09-710-339-2

Perfect score: 2684

Sequence: 1 MVAMWSLFLYGLQVADPALA.....LPRVLYPTEKLAGSKICSSS 498

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2684	100.0	498	22	AAB84206
2	2572	95.8	478	16	AA872450
3	2572	95.8	478	16	AA878270
4	2572	95.8	478	17	AA878270
5	2552	95.1	478	15	AA846005
6	2503.5	93.3	478	16	AA879025
7	1666.5	62.1	493	17	AA888212
8	1449	54.0	512	11	AA807574
9	1400.5	52.2	511	9	AA81180
10	1398.5	52.1	511	9	AA81161
11	1397	52.0	468	13	AA824136

12	1396.5	52.0	494	8	AA870571	Alpha-amylase gene
13	1392	51.9	468	15	AA863184	Variant alpha amyl
14	1391	51.8	468	15	AA863185	Variant alpha amyl
15	1387	51.7	468	15	AA863187	Variant alpha amyl
16	1338	49.9	452	15	AA863186	Variant alpha amyl
17	453	16.9	719	22	AA874220	Bacillus maltogeni
18	452.5	16.9	712	12	AA810051	Cyclomaltodextrin
19	452	16.8	719	22	AA874219	Bacillus maltogeni
20	450	16.8	719	20	AA870621	Bacillus maltogeni
21	450	16.8	719	20	AA870621	Amino acid sequenc
22	450	16.8	719	21	AA871731	Bacillus sp. malo
23	450	16.8	719	22	AA859956	Bacillus maltogeni
24	450	16.8	719	22	AA874216	Bacillus maltogeni
25	449.5	16.7	656	11	AA806109	Sequence of cyclom
26	446	16.6	719	22	AA874217	Bacillus maltogeni
27	446	16.6	719	22	AA874218	Bacillus maltogeni
28	446	16.6	719	22	AA874222	Bacillus maltogeni
29	446	16.6	719	22	AA874223	Bacillus maltogeni
30	444.5	16.6	713	11	AA806110	Sequence of cyclom
31	444.5	16.6	713	12	AA810052	Cyclomaltodextrin
32	444	16.5	719	22	AA874221	Bacillus maltogeni
33	432.5	16.1	683	17	AA817636	Thermoanaerobacter
34	432	16.1	682	17	AA817637	Thermoanaerobacter
35	430.5	16.0	683	17	AA817640	Thermoanaerobacter
36	428.5	16.0	655	13	AA820139	Sequence encoded b
37	427.5	15.9	681	17	AA817615	Thermoanaerobacter
38	427.5	15.9	683	17	AA817641	Thermoanaerobacter
39	427	15.9	526	19	AA844845	Cyclomaltodextrin
40	426.5	15.9	683	17	AA806772	Wild type cyclomal
41	426.5	15.9	683	17	AA817646	CGTase variant 146
42	426.5	15.9	683	17	AA817647	CGTase variant 145
43	426.5	15.9	683	17	AA817648	CGTase variant 145
44	426.5	15.9	683	17	AA817651	CGTase variant 145
45	426.5	15.9	683	17	AA817622	Thermoanaerobacter

ALIGNMENTS

RESULT 1

ID AAB84206 standard; Protein: 498 AA.

XX

AC AAB84206;

XX

DF 06-AUG-2001 (first entry)

XX

DE Amino acid sequence of a fungamyl-like alpha-amylase.

XX

KW Fungamyl-like alpha-amylase; glucoamylase; dextrinisation; maltose;

KW alcohol; starch; dough improver; brewing; starch liquification.

XX

OS Aspergillus oryzae.

XX

XX

PN WO200134784-A1.

XX

PD 17-MAY-2001.

XX

PE 10-NOV-2000; 2000WO-DK00626.

XX

PR 10-NOV-1999; 99DK-0001617.

XX

PA (NOVO) NOVOZYMES AS.

XX

PI Bisgard-Frantzen H, Svendsen A, Pedersen S;

XX

DR WPI: 2001-367478/38.

XX

XX N-PSDB; AAF90208.

PT New variant of Fungamyl-like alpha-amylase, useful for production of

PT maltose syrups, includes mutations that improve stability against heat

PT and acidic pH

XX

PS Claim 1: Page 42-45; 49pp; English.

CC The present sequence represents a fungamyl-like alpha-amylase. The
CC specification describes variants of this fungamyl-like alpha-amylase,
CC which have an alteration in one the amino acid regions 98-110, 150-160,
CC 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion
CC or substitution of an amino acid or an insertion of an amino acid
CC downstream of a particular position. The variants retain alpha-amylase
CC activity, and have better heat stability and/or stability at acidic pH,
CC relative to wild-type enzyme. The variants can therefore be used at
CC higher temperatures (more efficient conversion or faster reaction, and
CC have reduced need for cooling and reduced risk of contamination). The
CC variants may also be used in conjunction with other enzymes,
CC particularly glucoamylase during dextrinisation. The variants are
CC used to produce syrups, particularly of high maltose content, or alcohol,
CC from starch; as dough improver for baked goods; in brewing, to increase
CC fermentability of the wort; and for liquefaction of starch.

SQ Sequence 498 AA;

Query Match 100.0%; Score 2684; DB 22; Length 498;
Best Local Similarity 100.0%; Pred. No. 8.3e-227;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MYAMMSLFLYGLQVAPALATPADMRSSQSYFLTDREARTDGTATCMTADQKYCG 60
DB 1 mYAMMSLFLYGLQVAPALATPADMRSSQSYFLTDREARTDGTATCMTADQKYCG 60
OY 61 TWOGIIDLKLDYIQMGFTAIWITPYTAQLPQTAYGDAYHGYWQDIYSLNENGTADDL 120
DB 61 twogiidlKldyIQmgftaiWitpytaqlPqtaygDayhgywqdiYslneNGTaddl 120
OY 121 KALSSALHERGMYLWADVYANHMGYDAGSSVDYSEKPESSODYFHPFCFQNEODTO 180
DB 121 kaIssalhergmylWadvyaNHmgYdagssvdySEkpfESSodYfhpfcfQneodtQ 180
OY 181 VEDCWLGDNTVSLPDLDTTKDYVKNEMWVWSLVSNYSIDGLRIDTYKHVKDWP6YN 240
DB 181 vedcwlGdntvslPdlDTTKdyVknemWvwsLvsNysIdglRIdtyKHvkdWp6YN 240
OY 241 KAAGVYCIQEVLDGDPATYCPYQWMDGYLNPPIYPLINAFKSTSGSKMDLYNMINTVK 300
DB 241 kaagvyCIqevldgDPatYcPyqWmdgyLnppiyPlInafKstsgSKmdlyNmIntvk 300
OY 301 SOCPDSTLIGTEVENHNDNRFASYNNDIALAKNVAFTIINDGIPITVAGOEQHYAGND 360
DB 301 soCPdStliGtEveNHndnRfAsynNdIaLaKnvaFtiIndgIpItvAGoeqHyagNd 360
OY 361 PANREATWLSGYPTDSELEKLIIASANAIRNVAISKDTGCVTKKMPYKDDTTIARRKGT 420
DB 361 panreAtwlsGyptDseLEklIIasanaIRnvaISkdtGcvTKkmpYkddTTiARRkgt 420
OY 421 DGSQIVTILSNKGASDSTYLSISGAGYTAGOQLTEVIGCTTVTVGSDGNVPEPAGLIP 480
DB 421 dgsQivTilSnkgASdSTyLSisGagYtaGoqlTEvIGctTVtvGsdGnvPePaGlIp 480
OY 481 RVLYPEXLAGSKICSS 498
DB 481 rvLyPeXlaGskICss 498

RESULT 2

ID AAR72450 standard; Protein; 478 AA.

XX AAR72450;

DT 01-DEC-1995 (first entry)

DE Aspergillus oryzae alpha amylase (mature protein).

XX Alpha amylase; variant; enzyme; detergent; additive; dishwashing;

KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
KW Bacillus stearothermophilus; dyeing; bleaching; scouring; textile;
KW thermostable.

XX Aspergillus oryzae.

XX WO9510603-A.

XX 20-APR-1995.

XX 05-OCT-1994; 94WO-DK00370.

XX 08-OCT-1993; 93DK-0001133.

XX 02-FEB-1994; 94DK-0000140.

XX (NOVO) NOVO-NORDISK AS.

XX Svendsen A, Thellersen M;
PI Van der zee P, Bisgardfrantzen H, Borchert T;
XX WPI; 1995-161790/21.

DR WPI; 1995-161790/21.

XX New Bacillus derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance

XX Disclosure; Page 75-76; 105pp; English.

XX Variant alpha amylase enzymes which have improved washing and/or
CC as detergent additives. The enzymes have one or more amino acid
CC residues added, deleted or substituted. The variants can also be
CC used for textile desizing prior to scouring, bleaching and dyeing.

CC The variants have improved thermostability, acid/alkaline stability,
CC low temperature optimum; pH optimum; higher hydrolysis velocity and
CC improved tolerance to other composition constituents, e.g. oxidation
CC agents.

XX Sequence 478 AA;

SQ Sequence 478 AA;
Query Match 95.8%; Score 2572; DB 16; Length 478;
Best Local Similarity 99.8%; Pred. No. 5.3e-217;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 21 ATPADMRSSQSYFLTDREARTDGTATCMTADQKYCGTWOGIIDLKLDYIQMGFTAI 80
DB 1 atPdMRssqsyfLtdrEartDgtatCmtadQkycgTwogiIdlKldyIqmgftai 80
OY 81 WITPYTAQLPQTAYGDAYHGYWQDIYSLNENGTADDLKALSSALHERGMYLWADVYA 140
DB 81 witpytaqlPqtaygDayhgywqdiYslneNGTaddlKaIssalhergmylWadvya 140
OY 141 NHMGYDAGSSVDYSEKPESSODYFHPFCFQNEODTOVEDCWLGDNTVSLPDLDTTK 200
DB 141 nhmgYdagssvdySEkpfESSodYfhpfcfQneodtQvedcwlGdntvslPdlDTtk 200
OY 201 DVYKNEMWVWSLVSNYSIDGLRIDTYKHVKDWP6YNKAAGVYCIQEVLDGDPATYC 260
DB 201 dvYknemWvwsLvsNysIdglRIdtyKHvkdWp6YNkaagvyCIqevldgDPatyc 260
OY 261 PYQWMDGYLNPPIYPLINAFKSTSGSKMDLYNMINTVKSDCPDSTLIGTEVENHNDN 320
DB 261 pyqWmdgyLnppiyPlInafKstsgSKmdlyNmIntvksdCPdStliGtEveNHndn 320
OY 321 FASYTNDIALAKNVAEFTIINDGIPITVAGOEQHYAGNDPANREATWLSGYPTDSELYK 380
DB 321 faSyTndIaLaKnvaEftiIndgIpItvAGoeqHyagNdpanreAtwlsGyptDseLyk 380
OY 381 LIASANAIRNVAISKDTGCVTKKMPYKDDTTIARRKGTGDSQIVTILSNKGASDSTY 440
DB 381 liasanaIRnvaISkdtGcvTKkmpYkddTTiARRkgtGdsQivTilSnkgASdSTy 440
OY 441 LSLGAGYTAGOQLTEVIGCTTVTVGSDGNVPEPAGLIPRVLYPTPKLAGSKICSS 498
DB 441 lslGagYtaGoqlTEvIGctTVtvGsdGnvPePaGlIPrvLyPtPKlaGskICss 498

Db 421 lsisgagytgqqltevigctvtvsgdgnvpvmagglprvlypteklagskicss 478

RESULT 3

AAR78270

ID AAR78270 standard; Protein; 478 AA.

XX AAR78270;

XX 17-JAN-1996 (first entry)

DE Aspergillus oryzae alpha amylase (mature protein).

KW Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;

KW starch; thermostable; methionine; Bacillus licheniformis;

KW Bacillus amyloliquefaciens; Bacillus stearothermophilus;

KW Aspergillus oryzae.

XX Aspergillus oryzae.

XX WO9521247-A1.

XX 10-AUG-1995.

XX 05-OCT-1994; 94MO-DK00371.

XX 02-FEB-1994; 94DK-0000141.

XX (NOVO) NOVO-NORDISK AS.

XX Marcher D, Nilsson TE, Pedersen HH, Toft AH;

XX WPI; 1995-283767/37.

PT Use of an oxidation stable alpha-amylase - for simultaneous desizing

PT and bleaching or scouring of fabrics contg. starch or starch derivs.

PS Disclosure; Page 25-26; 37pp; English.

XX Oxidation stable alpha amylases can be used for the simultaneous

XX desizing and bleaching or scouring of a fabric comprising starch or

XX starch derivatives. They exhibit a better heat stability,

XX especially in the presence of oxidising agents. They are obtained

XX from a parent alpha amylase by replacing one or more methionine

XX residues with any amino acid different from Cys or Met, preferably

XX Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is

XX pref. derived from a Bacillus species, although alpha amylases of

XX fungal origin can also be used. This sequence is the wild type

XX (unmodified) alpha amylase of Aspergillus oryzae.

XX Sequence 478 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

OY 261 PYQNMVMDGLVNPYIPYPLLNAAFKSTSGSMDDLYNMINTVKSDCEDSTLLGTFVENHNDNR 320

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 241 PYQNMVMDGLVNPYIPYPLLNAAFKSTSGSMDDLYNMINTVKSDCEDSTLLGTFVENHNDNR 300

OY 321 FASNTNDIALAKNVAFAFIILNDGIPTIYAGOEHOYHAGNDPANEPATLSCGYPTDSELYK 380

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 301 FASYNDIALAKNVAFAFIILNDGIPTIYAGOEHOYHAGNDPANEPATLSCGYPTDSELYK 360

OY 381 LIASANAIRNVAISKDGTGFTYKKNPIYKDDTTIARRKGTGDSQIVTILSNKKGASGDSYT 440

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 361 LIASANAIRNVAISKDGTGFTYKKNPIYKDDTTIARRKGTGDSQIVTILSNKKGASGDSYT 420

OY 441 LLSGAGYTAGGOLTEVIGCTVTVMGSDGNVPVPMAGGLPRVLYPTTEKLASKICSS 498

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 421 lsisgagytgqqltevigctvtvsgdgnvpvmagglprvlypteklagskicss 478

RESULT 4

AAW14500

ID AAW14500 standard; Protein; 478 AA.

XX AAW14500;

XX 04-JUN-1997 (first entry)

DE Aspergillus oryzae alpha-amylase (mature protein).

XX alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;

XX Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;

XX calcium dependency; substrate binding; stability; pH optimum;

XX thermostability; cleavage; oligosaccharide substrate; dishwashing;

XX washing; detergent additive; fabric desizing; starch liquefaction;

XX sweetener; ethanol production; variant.

XX Aspergillus oryzae.

XX

XX

Key Location/Qualifiers

FT MISC-difference 13..45

FT /label= "loop 1 modification region

FT /note= "at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 7-23 of AAW14499 is deleted or

FT replaced with a fragment corresponding to

FT this fragment; claim 33"

FT MISC-difference 14..40

FT /label= "loop 1 modification region

FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a

FT template for a variant) corresponding to 8-18

FT of AAW14499 is deleted or replaced with a

FT fragment corresponding to this fragment; claim

FT 35" MISC-difference 28..42

FT /label= "loop 1 modification region

FT /note= "at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 12-19 of AAW14499 is deleted or

FT replaced with a fragment corresponding to

FT this fragment; claim 30"

FT MISC-difference 32..38

FT /label= "loop 1 modification region

FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a

FT template for a variant) corresponding to 14-15

FT of AAW14499 is deleted or replaced with a

FT fragment corresponding to this fragment; claim

FT 32" MISC-difference 66..84

FT /label= "loop 2 modification region

FT /note= "at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 44-57 of AAW14499 is deleted or

FT replaced with a fragment corresponding to

FT this fragment; claim 18"

FT MISC-difference 70..78


```

XX DE Mutant alpha-amylase.
XX KM Methionine substitution; stability; activity; detergent;
XX KW dishwashing agents; liquefaction agents.
XX OS Aspergillus oryzae.
XX PN WO9402597-A.
XX PD 03-FEB-1994.
XX PF 06-JUL-1993; 93WO-DK00230.
XX PR 23-JUL-1992; 92DK-0000946.
XX PR 16-DEC-1992; 92DK-0001503.
XX PR 15-MAR-1993; 93DK-0000292.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Bisgard-frantzen H, Svendsen A;
XX DR WPI; 1994-048855/06.
XX PT Mutant alpha-amylase from Bacillus species comprising a
XX PT methionine substitution - with improved stability and activity at
XX PT low pH, for use in detergents, dishwashing agents and
XX PT liquefaction agents
XX PS Claim 1; Page 7; 20pp; English.
CC CC The sequence os that of the Asoergillus oryzae alpha amylase, sold
CC CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC CC be mutated by substitution of one or more of its methionine residues
CC CC for any amino acid other than cysteine. The mutant alpha-amylase
CC CC exhibits a better activity level and better stability in the
CC CC presence of oxidising agents than previous mutant alpha amylases,
CC CC and improved thermostability at moderately low pH. The enzyme can
CC CC be used as an additive for detergents, dishwashing agents and
CC CC liquefaction agents.
SQ Sequence 478 AA;

Query Match 95.1%; Score 2552; DB 15; Length 478;
Best Local Similarity 99.4%; Pred. No. 3e-215;
Matches 475; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 21 ATPADMRQSISYFLITDRFARTDSTTATCNTADOKYCGGTWOGIIDKLDYIOGMGFTAI 80
DB 1 atpadmrqsisyflitdrfartdsgstatacntadqkycggtwgilidkldyiogmgtai 60
OY 81 WITPVTAQLPOTTAYGADYHGYWODIYSLNENYGTADDLKALSSALHERGMYLWVDVYA 140
DB 61 wltpvtaqlpottaygadyhgywqddylslnenygtadddlalssalhergmylmvdvya 120
OY 141 NHMGTDGAGSSVDVSVEKFPSSQDYFHFPCFIQNVEDQTOVEDCWLGDNTVSLPDLDTTK 200
DB 121 nhmgtdgagssvdvsvekfpsdqdyfhpfcfiqnvedqtovedcqlgdntvslpdltdtk 180
OY 201 DUVKNKWDWYGSIVSNSTIGLRIDTVKHQKDFWFGYKKAACYTCGVEYLDGDDPAYTC 260
DB 181 duvknkwdwysivsnstidgllridtvkhvqkdfwpgynaakayvcigevldgddpaytc 240
OY 261 PYONVMDGVNLPYIYPLINAFKSTSGSMDLYNNINVTVKSDCPDSTLLGTFVNHHPNPR 320
DB 241 pyonvmdgvnlpyiylplinafkstsgsmdllynnintvscdpcdstllgtfvenhnp 300
OY 321 FASTYNDIALKKNVAAFIILNDGIPIIYAGGEQHYAGGNDPANREATWLSGYPTDSELYK 380
DB 301 fastyndialaknvaafillndgipiliyaggeqhyaggnpanreatwlsgyptdselyk 360
OY 381 LIASANAIRNAISKDGFVTVYKKNMPLYKDDTTIAMKKGTGDSQIVITILSKGASGDSYR 440

```

```

DB 361 liasanairnaiskdgvfvyknmplykddittlamkxgdcgsivvtilnknkasgdsyt 420
OY 441 LSLSGAGTAAQQLTEVIGCTTWTGSDGNVPVPMAGLPVRVLYPTREKLAKGSKICSS 498
DB 421 lslegagyaagqltevigcttvtvgsdgnvppmagglprvlypteklagskicss 478

RESULT 6
AAR79025
ID AAR79025 standard; protein; 478 AA.
XX
XX AAR79025;
AC
XX 22-MAR-1996 (first entry)
DT
XX
XX Mature taka-amylase A.
DE
XX Wild type; neopullulanase; B. stearotheophilus; mutant;
KW food industry; modification; hydrophobicity; replacement
KW insertion; deletion.
XX
XX Aspergillus oryzae.
OS
XX
XX Key Location/Qualifiers
FH Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
XX JF07177891-A.
PN
XX
XX 18-JUL-1995.
PD
XX
XX 31-OCT-1994; 94JP-0288658.
PF
XX
XX 12-NOV-1993; 93JP-0306096.
PR
XX
XX (EZAK ) EZAKI GLICO CO.
PA
XX (NIDE ) NEC CORP.
PA
XX
XX WPI; 1995-279919/37.
DR
XX
XX
XX Modifying a transferase by enhancing hydrophobicity of a selected
XX PT site - increases transfer activity, also new mutant
XX PT neo-pullulanase(s)
XX
XX Disclosure; Page 10-11; 18pp; English.
XX
XX This sequence represents the mature form of taka-amylase from A. oryzae.
XX CC This sequence was used in a method for the generation of mutant
XX CC pullulanases for use in the food industry (see also AAR79026-28). The
XX CC wild type pullulanase enzyme was modified by the method of the
XX CC invention for enhancing the hydrophobicity of a selected site of the
XX CC pullulanase. The method comprises replacement of a group in the
XX CC selected site with a hydrophobic group, replacement of an amino acid
XX CC with a hydrophobic amino acid, and/or insertion or deletion of a
XX CC hydrophobic amino acid from the selected site. The method was used
XX CC to produce neopullulanases T377F, S422V and M375L.
XX
SQ Sequence 478 AA;

Query Match 93.3%; Score 2503.5; DB 16; Length 478;
Best Local Similarity 97.9%; Pred. No. 5.3e-211;
Matches 468; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

OY 21 ATPADMRQSISYFLITDRFARTDSTTATCNTADOKYCGGTWOGIIDKLDYIOGMGFTAI 80
DB 1 atpadmrqsisyflitdrfartdsgstatacntadqkycggtwgilidkldyiogmgtai 60
OY 81 WITPVTAQLPOTTAYGADYHGYWODIYSLNENYGTADDLKALSSALHERGMYLWVDVYA 140

```

```

Db      61  wltptvaqlpqdcaygdaytgywtldtlyslnenygtadddlkalasalthergmjlmvav 120
QY      141  NHMGTIDAGSSVDYSVFEPFSSQDYFHFPCFIQNYEDQTOVEDCWLGDNTVSLPDLDTTK 200
Db      121  nhmgydgagssvdysvfikpfssqdyfhpfcfiqnyedqtdyvedcwlqdnltvslpdlctk 180
QY      201  DVKKKEWDMWGSILVSNYSIDGLRIDTVKHKVQKDFWPGYNKAKGVYCCGEVLDDGPAYTC 260
Db      181  dvkkewdmwgsilvsnysidglridtvkhqkdfwpgynkaagyccgevlddgdpaytc 240
QY      261  PYQNVMDGVLYNPPIYPLLNFKSTSGSMDLLYMKINTVKSDCPDSTLLGTVEVHNHNP 320
Db      241  pyqnmvdgvllynpilynaftksqsmddlylmintvksdcpdstllgtfvenhnp 300
QY      321  FASYNNDIALAKNVAFTIINDGPIITVYAGQOHYAGGNDPANREATWLSGYPDSELYK 380
Db      301  fasyndialaknvaftiindgpiityaqeqhyaggnpanreatwlsqypdselyk 360
QY      381  LIASANAIRNVAISKDTGFVYKKNPFIKDDTTIAMRKGTGSOIVTILSNKASGDSYT 440
Db      361  liasanaairnvaiskdtgfvykn-pykddttiamrtgtdsqdvtllsnkagsgdslt 419
QY      441  ILSISAGYTAAGQOLEVIGCTTVTVGSDGNVPVPMAGGLPRVLYPTEKLAGSKICSS 498
Db      420  lsisagytgaqglevlgcttvsvgsdgnvpvpmagglprvlypteklagskicsds 477

```

```

RESULT 7
AAR88212
ID      AAR88212 standard; Protein: 493 AA.
XX
AC      AAR88212;
XX
DT      03-APR-1996 (first entry)
XX
DE      Alpha-amylase.
XX
KW      Alpha-amylase; thermostable enzyme; baking; Aspergillus niger.
OS      Thermomyces lanuginosus CBS 224.63.
XX
FH      Key
FT      Location/Qualifiers
ET      1..18
PT      Peptide
XX
XX      W09601323-A1.
XX      18-JAN-1996.
XX      03-JUL-1995; 95WO-EP02607.
XX      04-JUL-1994; 94GB-0013419.
XX      (DANI-) DANISCO AS.
XX      Michelsen B, Rasmussen P;
XX      MPI: 1996-087673/09.
XX      N-PSDB; AAT10562.
XX
XX      Thermophilic alpha-amylase with activity range of 60-80 degrees C -
PT      derived from Thermomyces lanuginosus, useful in the prepn. of
PT      foodstuffs and bakery prods. esp. bread
XX
PS      Claim 3: Page 36-38; 94pp; English.
XX
XX      A thermostable alpha-amylase (AAR88212) of Thermomyces lanuginosus
CC      CBS 224.63 is expressed in Aspergillus niger hosts using a gene
CC      (AAT10562) isolated from a T. lanuginosus gene library. The
CC      recombinant enzyme (54-60 kDa) shows optimal activity at
CC      60-70 deg and pH 5.8-6, has a pI of 3.7 and is active at 60-80
CC      deg.
XX
XX

```

SQ Sequence 493 AA:

```

Query Match      62.1%; Score 1666.5; DB 17; Length 493;
Best Local Similarity 63.2%; Pred. No. 17e-137;
Matches 308; Conservative 69; Mismatches 101; Indels 9; Gaps 4;

QY      16  APAL--AAPADRRSOSIYFLTRFAPRDGSTTATCNTRADKRCGCGTWQGIIDRLDYIO 73
Db      12  spllyraalpdeakagslyfmldtrfardcnstapcdctakkygsgvrgjinnldyiq 71
QY      74  GMGFTAIWTPVTAQLPQTTAVGDA--YHGWOQDIYSINENYGTADLKAISALHENG 131
Db      72  dmgtfaiwtpvtaqwdvdaadaatsyngyqkdlylsnksfgytaddlkaladahlary 131
QY      132  MYLMDVYVNNHMGYDAGSSVDYSVFEPFSSQDYFHFPCFIQNYEDQTOVEDCWLGDNTV 191
Db      132  mlmvdvvanhfyggsdshvdytlnpnsqdyfhpfcfiedydngevegcwladpct 191
QY      192  SLPDLPTKDVKNENYDMWGSILVSNYSIDGLRIDTVKHKVQKDFWPGYNKAKGVYCCIGEY 251
Db      192  tlpdvtltpqvrtffndwkslvansyidglrvctkhvekdfwpdfneaac-tvgev 250
QY      252  LDGDPAYTCFYQNVMDGVLYNPPIYPLLNFKSTSGSMDLLYMKINTVYKSDCPDSTLLGT 311
Db      251  lngdpaytcpygvlqgvllynpilypaldafkvgnglgljaqaltvgesckdsnllgn 310
QY      312  FVENHNDPRFASYNNDIALAKNVAFTIINDGPIITVYAGQOHYAGGNDPANREATWLSG 371
Db      311  flenhndprfasyddialaknglatilisdgipilytqeqhyagdnhpnreawvlsq 370
QY      372  YPTDSELYKLIASANAIRNVAISKDTGFVYKKNPFIKDDTTIAMRKGTGSOIVTILSN 431
Db      371  yntdaelygfikangirnaisqneftsckkkyvgdstdalnrg---gvvvlsn 426
QY      432  KGASGDSYTLISGAGYTAAGQOLEVIGCTTVTVGSDGNVPVPMAGGLPRVLYPTEKLAG 491
Db      427  egaagelvtvslpqtgfaegteltvviscktlvagsdgvdpysglpsvlypsqjak 486
QY      492  SKICSS 498
Db      487  sqlcasa 493

RESULT 8
AAR07574
ID      AAR07574 standard; protein; 512 AA.
XX
AC      AAR07574;
XX
DT      02-FEB-1991 (first entry)
XX
DE      Alpha-amylase encoded by the AMY1 gene.
XX
XX      Schwanniomycetes yeast cells; expression cassette; alpha-amylase;
KW      marker gene; region; signal peptide; terminator;
KW      autonomously replicating sequence.
XX
XX      Schwanniomycetes occidentalis.
XX
XX      Key
FH      Location/Qualifiers
FT      Disulfide-bond 66..74
FT      /note="1"
FT      Disulfide-bond 186..200
FT      /note="2"
FT      Disulfide-bond 276..319
FT      /note="3"
FT      Disulfide-bond 475..510
FT      /note="4"
XX
XX      EP394538-A.
XX      31-OCT-1990.
XX
XX

```



```

XX 28-APR-1989; 89EP-0107780.
XX
XX 28-APR-1989; 89EP-0107780.
XX
XX (RHEI ) RHEIN BIOTECH GES.
XX
XX Hollenberg C, Strasser A;
XX
XX WPI; 1990-328670/44.
XX
XX N-PSDB; AAQ06388.
XX
XX Transformed Schwanniomycos yeast cells - contg. an expression
XX cassette contg. regulon, DNA coding for foreign protein and
XX terminator
XX
XX PS Disclosure; Fig 1B(1-6); 59pp; English.
XX
XX The expression cassette may contain a regulon (R), a signal peptide
XX sequence (S), a foreign protein sequence (F) and a terminator (T)
XX of which R and/or S and/or T are derived from the alpha-amylase gene
XX of S. occidentalis.
XX R is pref. a 1.8 kb BglII-XhoI fragment (bases -1 to -540 of this
XX sequence).
XX S contains all or part of one of the sequences represented
XX in AAQ06383-87.
XX F is e.g. a cellulase, intereukin, insulin-like-growth factor,
XX interferon etc.
XX T is pref. all or part of the terminator of this sequence
XX (bp 1537-1740).
XX AMY1 acts as selective marker for the vector carrying the expression
XX cassette.
XX See also AAQ06389.
XX
XX Sequence 512 AA;

```

```

Query Match 54.0%; Score 1449; DB 11; Length 512;
Best Local Similarity 56.4%; Pred. No. 2.3e-118;
Matches 269; Conservative 74; Mismatches 132; Indels 2; Gaps 2;

```

```

QY 21 ATPAD-WRSOSITYFLDREARFRTGSGTTATCNTADOKYCGGTGOGIIDLKIDYIQMGFTFA 79
   :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 36 sssadkwwdgslyivtrfarsdgtstacdclvsdrkycgsgyqjldkldyiqmgfta 95
QY 80 IWTPPVTAQLPQTAYAGDAGYAGWQODIYSLNENYGTADLKLSSALHRCGYLWADVY 139
   ||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 96 lwspvveqipdnatayayayngymkndelntnftgadclqjaaselharsmlmadvv 155
QY 140 ANHMGYDGAAGSSVDYSVKFRPSSODYFRHPCFIONYEDQYQVEDCWMGDMTVSLPDLDT 199
   ||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 156 ynhyawngdssvdsftrfngqsyfhdyclfthnyndqtnvedcwegdevslpdlste 215
QY 200 KDVENENYDWVGSILVSNYSIDGLRIDTVKHQKDFWPGYNKAAGVYCIIEVDLGPAPYT 259
   ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 216 dneviygvfqtwsdvtqysidqrlrdsaknvdtaasltkfedasgyynlgevygqdpct 275
QY 260 CPYQNVMDGVLTNPYIYPLNAFKSTSGSMDLXNMINTVKSDDCPDSTLLIGTVEVNDNP 319
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 276 cpyqymkgyvtnplypyvyrffsdatsatseletsmistlqsscsdvsllgnfienhdqy 335
QY 320 RFASTYNIDIALAKNVAFTIILNDGPIITYAGQEOHYAGNDPANREATWLSGYPDSELY 379
   ||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 336 rtpsvtstdtslikndmafiilgdqilpilyyqegqllngsdpanrealwlsygnldeqy 395
QY 380 KLIASAMIRNVAISKDTEVYTKNMPYIKDPTIAMRKGTGSOIVTILSNKGASGDSY 439
   :||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 396 elisklmgqlrqaikkdasytkysvssedhyiactrksdangqlisltfnlgsng-sq 454
QY 440 TSLSGAGTAAQQLTEVIGCTVYAGSDGNVPMAGMGPRVLYPREKLAGAGTICS 496
   :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 455 dlvtentcyssgdkiydliscnsvlaagdsisvsiisgmppayapasvslsgsich 511

```

RESULT 9

```

AAP81180
ID AAP81180 standard; protein; 511 AA.
XX
XX AAP81180;
AC
XX 12-DEC-1990 (first entry)
DE
XX Sequence of alpha-amylase encoded by genomic DNA of Schwanniomycos
DE castellii.
XX
XX Brewing; beer; breadmaking; biomass.
XX
XX OS Schwanniomycos castellii.
XX
XX FN EP260404-A.
XX
XX PD 23-MAR-1988.
XX
XX PF 17-JUL-1987; 87EP-0110370.
XX
XX PR 17-JUL-1987; 87EP-0110370.
XX
XX PA (HETB ) HEINEKEN TEC BEHEER NV.
XX
XX PI Strasser A, Martens FB, Dohmen J, Hollenberg CP;
XX
XX DR WPI; 1988-078794/12.
XX
XX N-PSDB; AAN81525.
XX
XX Polypeptide prodn. from cells transformed with yeast DNA -
XX esp. coding alpha amylase or glucoamylase, able to convert
XX starch to ethanol
XX
XX PS Claim 18; Fig 2A-2C; 74pp; English.
XX
XX A cosmid library is constructed from genomic DNA of ATCC26076 and this
XX screened for inserts contg. the AA gene by ability to transform
XX S.cerevisiae GRF 18 to histidine prototrophy. One positive cosmid
XX (33kb; pyc1) was digested with EcoRI, religated and used to transform
XX E.coli JAE21. Positive transformants contain pyc1-alpha, which includes
XX a 5kb EcoRI fragment contg. the gene in AAN81525.
XX
XX Sequence 511 AA;

```

```

Query Match 52.2%; Score 1400.5; DB 9; Length 511;
Best Local Similarity 55.1%; Pred. No. 4.1e-114;
Matches 264; Conservative 74; Mismatches 134; Indels 7; Gaps 4;

```

```

QY 21 ATPAD-WRSOSITYFLDREARFRTDGSSTTANCNTADOKYCGGTGOGIIDLKIDYIQMGFTFA 79
   :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 36 sssadkwwdgslyivtrfarsdgtstacdclvsdrkycgsgyqjldkldyiqmgfta 95
QY 80 IWTPPVTAQLPQTAYAGDAGYAGWQODIYSLNENYGTADLKLSSALHRCGYLWADVY 137
   ||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 96 lwspvveqipdnatayayayngymkndelntnftgadclqjaaselharsmlmadvv 152
QY 140 ANHMGYDGAAGSSVDYSVKFRPSSODYFRHPCFIONYEDQYQVEDCWMGDMTVSLPDLDT 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 156 ynhyawngdssvdsftrfngqsyfhdyclfthnyndqtnvedcwegdevslpdlste 212
QY 198 IWKDVYKNEWMDWVGSILVSNYSIDGLRIDTVKHQKDFWPGYNKAAGVYCIIEVDLGPAPYT 257
   ||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 213 dneviygvfqtwsdvtqysidqrlrdsaknvdtaasltkfedasgyynlgevygqdpct 272
QY 258 YTCPYQNVMDGVLTNPYIYPLNAFKSTSGSMDLXNMINTVKSDDCPDSTLLIGTVEVNDNP 317
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 273 ycpqymkgyvtnplypyvyrffsdatsatseletsmistlqsscsdvsllgnfienhd 332
QY 318 NREFASTYNIDIALAKNVAFTIILNDGPIITYAGQEOHYAGNDPANREATWLSGYPDSE 377
   ||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


Db 362 LIAKANAARNAAYQDSSYATSGISVIFSNHVIATKRGs---VVSVFNLGSSGSS-d 416
QY 441 LSLGAGTTAGQQLTEVIGCTTWTGSDGNVPYPMAGLPRVLVPTREKLASGKICS 496
Db 417 VLIANTGYSSEGLVEVLTCSTVSGSSD--LQVSIQGGPQLIVPAK--YASDLICS 468

RESULT 15

AAR63187 standard; Protein; 468 AA.

AAR63187:

16-JUN-1995 (first entry)

Variant alpha amylase deriv. from Saccharomyopsis fibuligera.

alpha amylase: carbohydrate hydrolase; increased activity;

tyrosine residue; enzyme centre; mass production; oligosaccharide;

variant; cyclomaltoextrin glucanotransferase.

Saccharomycopsis fibuligera.

Key Location/Qualifiers

JP06253836-A.

13-SEP-1994.

04-MAR-1993: 93JP-00693303.

04-MAR-1993: 93JP-00693303.

(AGEN) AGENCY OF IND SCI & TECHNOLOGY.

WP1: 1994-328987/41.

N-PSDB; AAQ71668.

Variant carbohydrate hydrolase(s) with increased activity -

consists of e.g. alpha-amylase with tyrosine residue in enzyme

centre, useful for mass-prodn. of oligosaccharide(s)

Example 1; Page 20-23; 27pp; Japanese.

AAR63184-7 show the amino acid sequences of variant alpha amylases, composed by substituting bases 247-249 of the structural gene (see AAQ71665-8) region, with TTC, TGC, CTC or AAC. These substitutions result in the 83rd amino acid residue (tyrosine) of the wild type sequence being changed to phenylalanine, tryptophan, leucine or asparagine respectively. The substituted amino acid is present in the active site of the enzyme and confers increased activity on the enzyme. The variants are useful for the mass production of oligosaccharides. (see AAR63188 for the variant structure of a cyclomaltoextrin glucanotransferase).

Sequence 468 AA;

Query Match 51.7%; Score 1387; DB 15; Length 468;
Best Local Similarity 54.4%; Pred. No. 5.5e-113;
Matches 259; Conservative 80; Mismatches 127; Indels 10; Gaps 5;

QY 22 TPAD-WRSOSTYFLTDREFARTDGTATCNTADOKYCGTMOGIIIDKDIYIGMGFTAI 80
Db 2 Tnadkwsrsqslqvtldfartdgtatscndetrlfcgsgtqllkldyikdmgtfai 61
QY 81 WITPVTALPQTAYGDAVHGYWODIYSLNENGTADLKAISALHERGMYLWVVA 140
Db 62 wispyvenipentatayayaybgmknkylkinenftaddlkslqelhdtdmlmndivt 121

QY 141 NHMGYDAGSSVDYSVFKPPSSODTFHPFCFIONTEDQTVEDCWLGDNTVSLPDLDTK 200
Db 122 nhysdsgsgsidyseylfndqkyfhnycilsnyddqavqsvcsweegdsvalpdlrted 181
QY 201 DVVKNEMYDWVGSLSVNSYSIDGLRIDFKHVKOKDFPGYNKAAGVYCIGEVLDGDPAYTC 260
Db 182 sdvasvfmswvkdftvgnysidglrldsakhvdqgffpdlfvsasgyysvgevfgdpaytc 241
QY 261 PYQNVMDGVNLNYPPIYPLLNAPKSTGSMDDLYNMTNTVKSDDCPDSTLLGTFVENHNDPR 320
Db 242 pyqnyipgvsnypllypdltrfktldssselqgmssvasscsdpcllntfvenhndher 301
QY 321 FASYTNDIALAKNYAARFIILNDGIPRIYAGQEQHAYAGNDPANREAPTWISGYPDSELYK 380
Db 302 fasmtdqslismaiafvlvgdipvlvy9qeqqlsgskdpmrrealwlsynkesdyk 361
QY 381 LIASANAIRNVAISKDGFVTYKNMPLYKDDTTIANKGTDSQIYVILSNKASGDSYT 440
Db 362 LIAKANAARNAAYQDSSYATSGISVIFSNHVIATKRGs---VVSVFNLGSSGSS-d 416
QY 441 LSLGAGTTAGQQLTEVIGCTTWTGSDGNVPYPMAGLPRVLVPTREKLASGKICS 496
Db 417 VLIANTGYSSEGLVEVLTCSTVSGSSD--LQVSIQGGPQLIVPAK--YASDLICS 468

Search completed: June 13, 2002, 08:53:25
Job time: 275 sec
